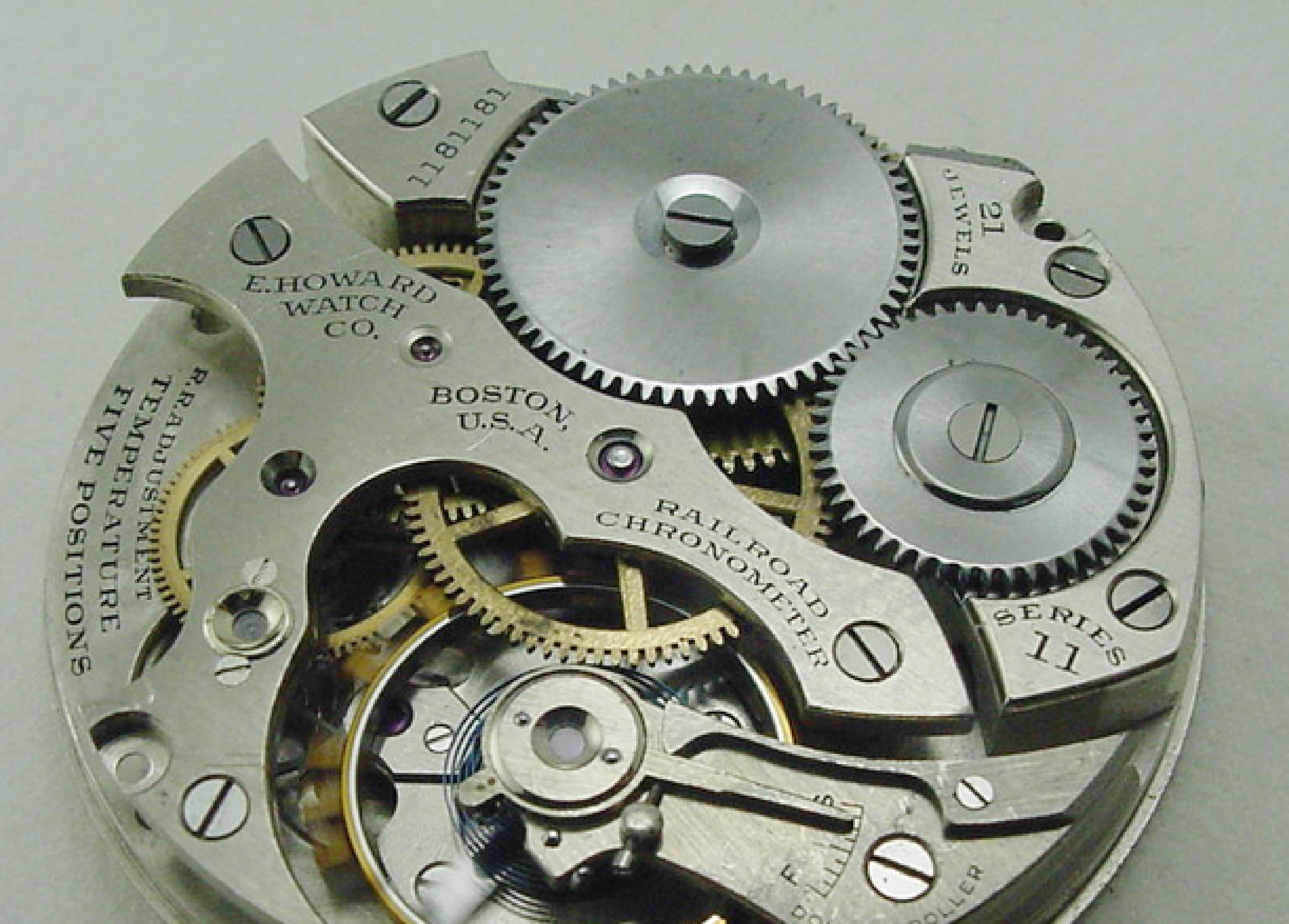


Spring Mechanics

Maurice Kendall

First Edition, 2012

ISBN 978-81-323-4539-8

Published by:
The English Press
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

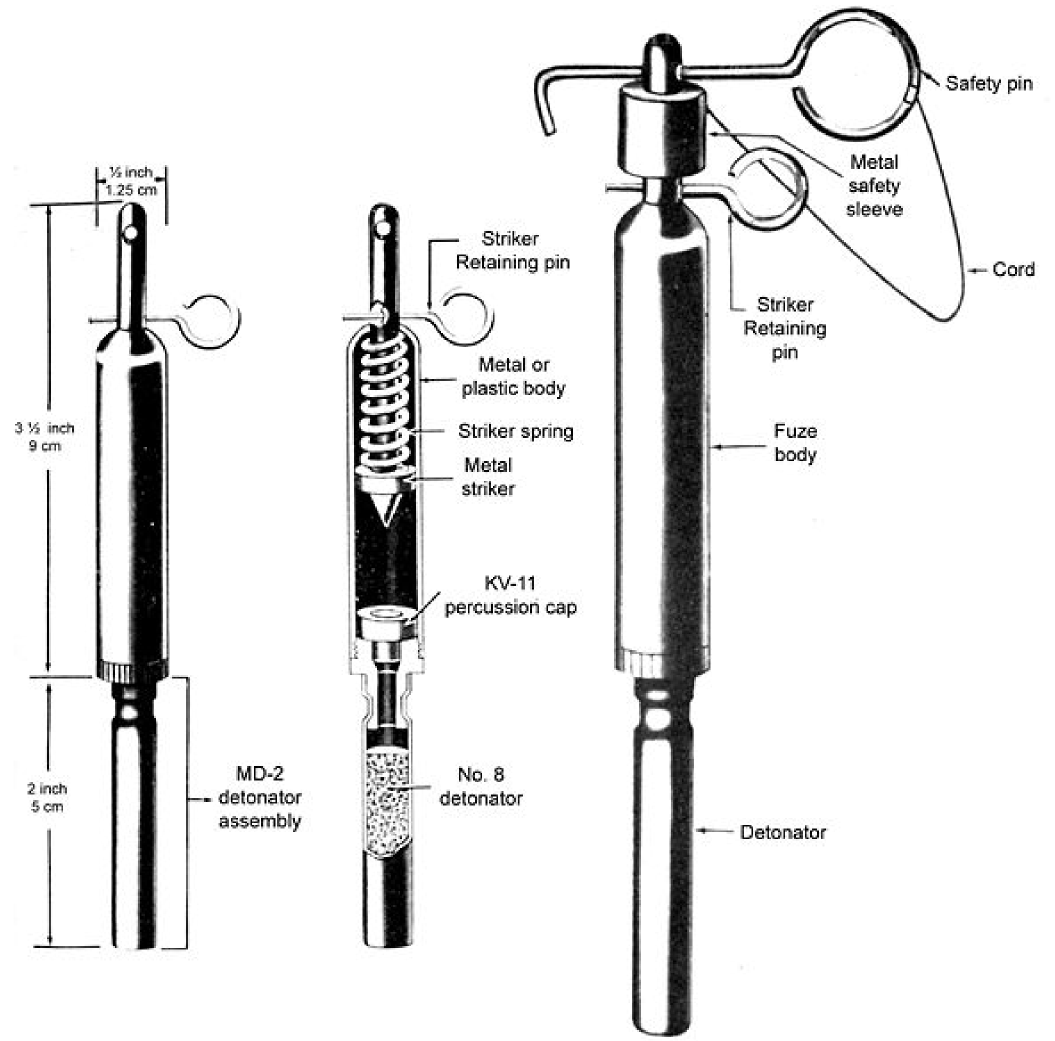
Chapter 1

Spring (Device)

Helical or *coil* springs designed for tension

Compression springs store energy when compressed

The English longbow - a simple but very powerful spring made of yew, measuring 2 m (6 ft 6 in) long, with a 470 N (105 lbf) draw force



Military boobytrap firing device from USSR (normally connected to a tripwire) showing spring-loaded firing pin

A **spring** is an elastic object used to store mechanical energy. Springs are usually made out of hardened steel. Small springs can be wound from pre-hardened stock, while larger ones are made from annealed steel and hardened after fabrication. Some non-ferrous metals are also used including phosphor bronze and titanium for parts requiring corrosion resistance and beryllium copper for springs carrying electrical current (because of its low electrical resistance).

When a spring is compressed or stretched, the force it exerts is proportional to its change in length. The *rate* or *spring constant* of a spring is the change in the force it exerts, divided by the change in deflection of the spring. That is, it is the gradient of the force versus deflection curve. An extension or compression spring has units of force divided by distance, for example lbf/in or N/m. Torsion springs have units of force multiplied by distance divided by angle, such as N·m/rad or ft·lbf/degree. The inverse of spring rate is compliance, that is: if a spring has a rate of 10 N/mm, it has a compliance of 0.1 mm/N.

The stiffness (or rate) of springs in parallel is additive, as is the compliance of springs in series.

Depending on the design and required operating environment, any material can be used to construct a spring, so long as the material has the required combination of rigidity and elasticity: technically, a wooden bow is a form of spring.

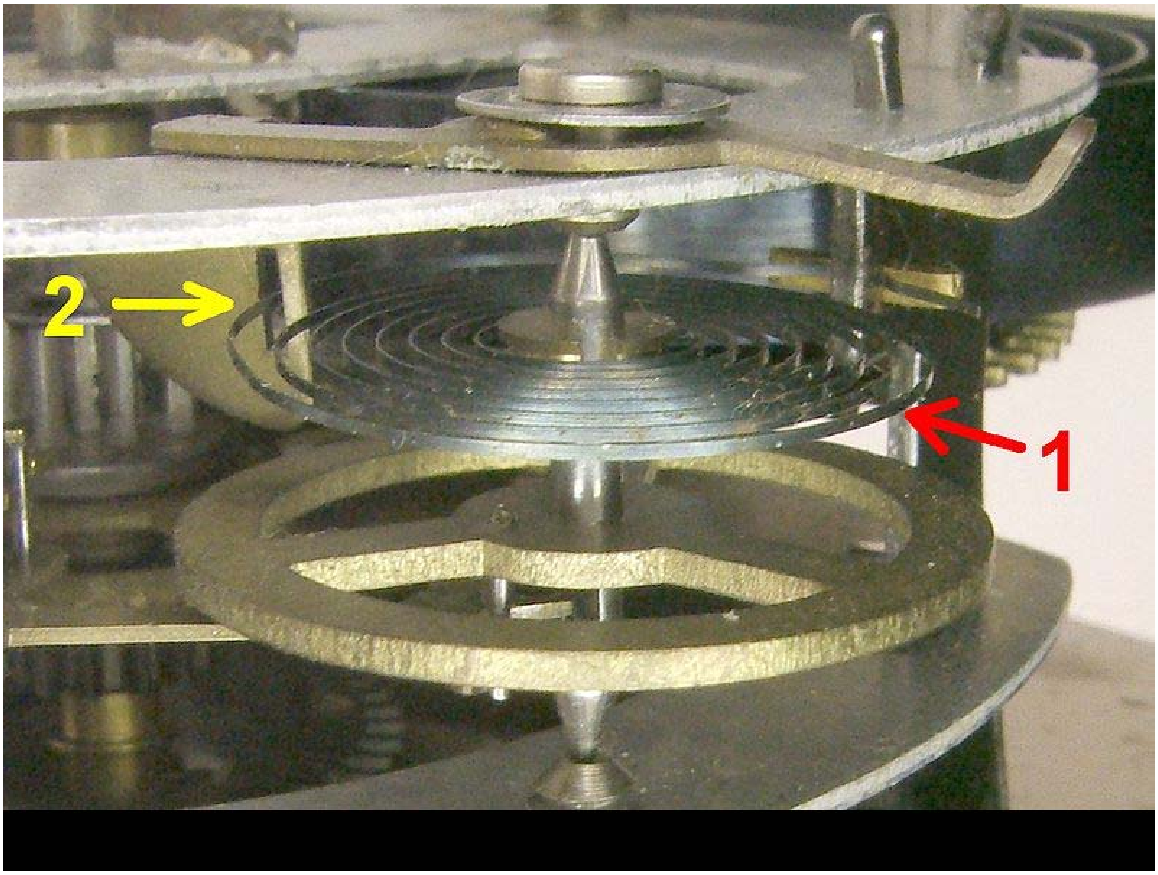
History

Simple non-coiled springs were used throughout human history e.g. the bow (and arrow). In the Bronze Age more sophisticated spring devices were used, as shown by the spread of tweezers in many cultures. Ctesibius of Alexandria developed a method for making bronze with spring-like characteristics by producing an alloy of bronze with an increased proportion of tin, and then hardening it by hammering after it is cast.

Coiled springs appeared early in the 15th century, in door locks. The first spring powered-clocks appeared in that century and evolved into the first large watches by the 16th century.

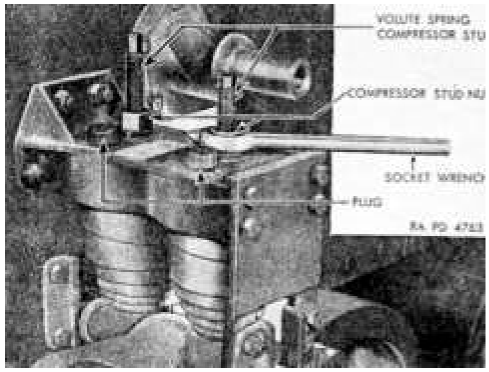
In 1676 British physicist Robert Hooke discovered the principle behind springs' action, that the force it exerts is proportional to its extension, now called Hooke's law.

Types

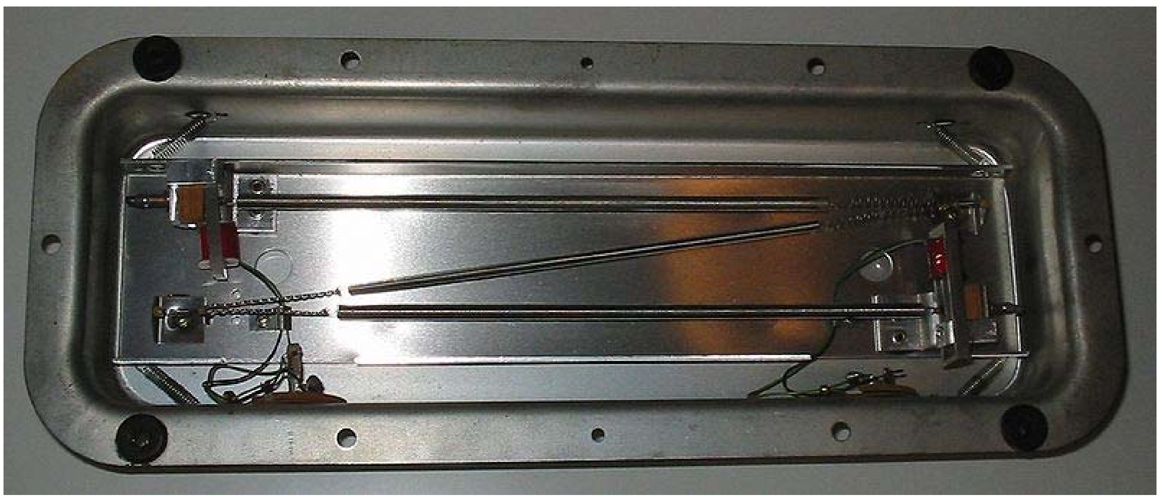


A spiral torsion spring, or hairspring, in an alarm clock.

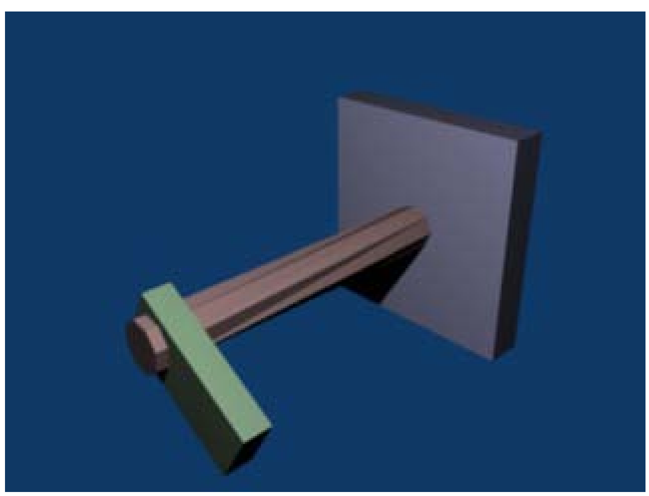
A volute spring. Under compression the coils slide over each other, so affording longer travel.



Vertical volute springs of Stuart tank



Tension springs in a folded line reverberation device.



A torsion bar twisted under load

Leaf spring on a truck

Springs can be classified depending on how the load force is applied to them:

- Tension/Extension spring - the spring is designed to operate with a tension load, so the spring stretches as the load is applied to it.
- Compression spring - is designed to operate with a compression load, so the spring gets shorter as the load is applied to it.
- Torsion spring - unlike the above types in which the load is an axial force, the load applied to a torsion spring is a torque or twisting force, and the end of the spring rotates through an angle as the load is applied.

They can also be classified based on their shape:

- Coil spring - this type is made of a coil or helix of wire
- Flat spring - this type is made of a flat or conical shaped piece of metal.

The most common types of spring are:

- Cantilever spring - a spring which is fixed only at one end.
- Coil spring or helical spring - a spring (made by winding a wire around a cylinder) and the conical spring - these are types of torsion spring, because the

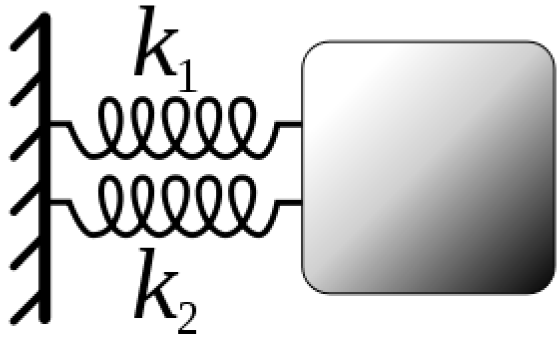
wire itself is twisted when the spring is compressed or stretched. These are in turn of two types:

  - *Compression springs* are designed to become shorter when loaded. Their turns (loops) are not touching in the unloaded position, and they need no attachment points.
    - A *volute spring* is a compression spring in the form of a cone, designed so that under compression the coils are not forced against each other, thus permitting longer travel.
  - *Tension* or *extension springs* are designed to become longer under load. Their turns (loops) are normally touching in the unloaded position, and they have a hook, eye or some other means of attachment at each end.
- Hairspring or balance spring - a delicate spiral torsion spring used in watches, galvanometers, and places where electricity must be carried to partially-rotating devices such as steering wheels without hindering the rotation.
- Leaf spring - a flat springy sheet, used in vehicle suspensions, electrical switches, bows.
- V-spring - used in antique firearm mechanisms such as the wheellock, flintlock and percussion cap locks.

Other types include:

- Belleville washer or Belleville spring - a disc shaped spring commonly used to apply tension to a bolt (and also in the initiation mechanism of pressure-activated landmines).
- Constant-force spring — a tightly rolled ribbon that exerts a nearly constant force as it is unrolled.
- Gas spring - a volume of gas which is compressed.
- Ideal Spring - the notional spring used in physics: it has no weight, mass, or damping losses.
- Mainspring - a spiral ribbon shaped spring used as a power source in watches, clocks, music boxes, windup toys, and mechanically powered flashlights
- Negator spring - a thin metal band slightly concave in cross-section. When coiled it adopts a flat cross-section but when unrolled it returns to its former curve, thus producing a constant force throughout the displacement and *negating* any tendency to re-wind. The commonest application is the retracting steel tape rule.
- Progressive rate coil springs - A coil spring with a variable rate, usually achieved by having unequal pitch so that as the spring is compressed one or more coils rests against its neighbour.
- Rubber band - a tension spring where energy is stored by stretching the material.
- Spring washer - used to apply a constant tensile force along the axis of a fastener.
- Torsion spring - any spring designed to be twisted rather than compressed or extended. Used in torsion bar vehicle suspension systems.
- Wave spring - a thin spring-washer into which waves have been pressed.

Physics



Two springs attached to a wall and a mass. In a situation like this, the two springs can be replaced by one with a spring constant of $k_{eq}=k_1+k_2$.

Hooke's law

Most springs (not stretched or compressed beyond the elastic limit) obey Hooke's law, which states that the force with which the spring pushes back is linearly proportional to the distance from its equilibrium length:

$$F = -kx,$$

where

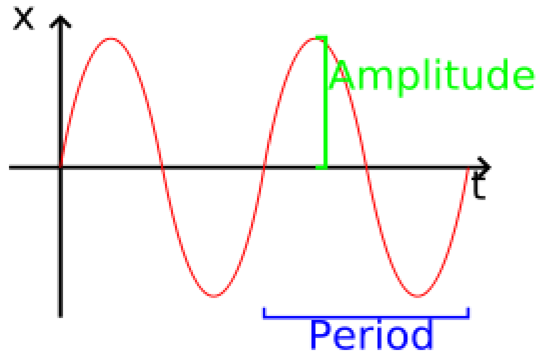
x is the displacement vector - the distance and direction in which the spring is deformed
F is the resulting force vector - the magnitude and direction of the restoring force the spring exerts
k is the **spring constant** or **force constant** of the spring.

Coil springs and other common springs typically obey Hooke's law. There are useful springs that don't: springs based on beam bending can for example produce forces that vary nonlinearly with displacement.

Simple harmonic motion

Since force is equal to mass, *m*, times acceleration, *a*, the force equation for a spring obeying Hooke's law looks like:

$$F = ma \quad \Rightarrow \quad -kx = ma.$$



The displacement, *x*, as a function of time. The amount of time that passes between peaks is called the period.

The mass of the spring is assumed small in comparison to the mass of the attached mass and is ignored. Since acceleration is simply the second derivative of x with respect to time,

$$-kx = m\frac{d^2x}{dt^2}.$$

This is a second order linear differential equation for the displacement *x* as a function of time. Rearranging:

$$\frac{d^2x}{dt^2} + \frac{k}{m}x = 0,$$

the solution of which is the sum of a sine and cosine:

$$x(t) = A\sin\left(t\sqrt{\frac{k}{m}}\right) + B\cos\left(t\sqrt{\frac{k}{m}}\right).$$

A and *B* are arbitrary constants that may be found by considering the initial displacement and velocity of the mass. The graph of this function with *B* = 0 (zero initial position with some positive initial velocity) is displayed in the image on the right.

Theory

In classical physics, a spring can be seen as a device that stores potential energy, specifically elastic potential energy, by straining the bonds between the atoms of an elastic material.

Hooke's law of elasticity states that the extension of an elastic rod (its distended length minus its relaxed length) is linearly proportional to its tension, the force used to stretch it.

Similarly, the contraction (negative extension) is proportional to the compression (negative tension).

This law actually holds only approximately, and only when the deformation (extension or contraction) is small compared to the rod's overall length. For deformations beyond the elastic limit, atomic bonds get broken or rearranged, and a spring may snap, buckle, or permanently deform. Many materials have no clearly defined elastic limit, and Hooke's law can not be meaningfully applied to these materials.

Hooke's law is a mathematical consequence of the fact that the potential energy of the rod is a minimum when it has its relaxed length. Any smooth function of one variable approximates a quadratic function when examined near enough to its minimum point; and therefore the force — which is the derivative of energy with respect to displacement — will approximate a linear function.

Force of fully compressed spring

$$F_{max} = \frac{Ed^4(L - nd)}{16(1 + \nu)(D - d)^3 n}$$

where

E - Young's modulus
d - spring wire diameter
L - free length of spring
n - number of active windings
ν - Poisson ratio
D - spring outer diameter

Zero-length springs

"Zero-length spring" is a term for a specially-designed coil spring that would exert zero force if it had zero length. That is, in a line graph of the spring's force versus its length, the line passes through the origin. Obviously a coil spring cannot contract to zero length because at some point the coils will touch each other and the spring will not be able to shorten any more. Zero length springs are made by manufacturing a coil spring with built-in tension, so if it *could* contract further, the equilibrium point of the spring, the point at which its restoring force is zero, occurs at a length of zero. In practice, zero length springs are made by combining a "negative length" spring, made with even more tension so its equilibrium point would be at a "negative" length, with a piece of inelastic material of the proper length so the zero force point would occur at zero length.

A spring with zero length can be attached to a mass on a hinged boom in such a way that the force on the mass is almost exactly balanced by the vertical component of the force from the spring, whatever the position of the boom. This creates a pendulum with very long period. Long-period pendulums enable seismometers to sense the slowest waves from earthquakes. The LaCoste suspension with zero-length springs is also used in gravimeters because it is very sensitive to changes in gravity. Springs for closing doors are often made to have roughly zero length so that they will exert force even when the door is almost closed, so it will close firmly.

Uses

- Balance springs mechanical timepieces
- Buckling spring keyboards
- In CD players
- Inside a pen
- Mattress
- Pogo Stick
- Slinky
- Trampoline
- Vehicle suspension

Chapter 2

Coil Spring and Springboard

Coil spring

A compression coil spring

A tension coil spring

Oxy-cut spring showing deformation due to loss of tempering in adjacent turn

A selection of conical coil springs

A **Coil spring**, also known as a *helical spring*, is a mechanical device, which is typically used to store energy and subsequently release it, to absorb shock, or to maintain a force between contacting surfaces. They are made of an elastic material formed into the shape of a helix which returns to its natural length when unloaded.

Coil springs are a special type of torsion spring: the material of the spring acts in torsion when the spring is compressed or extended.

Metal coil springs are made by winding a wire around a shaped former - a cylinder is used to form cylindrical coil springs.

Variants

The two usual types of coil spring are:

- Tension coil springs, designed to resist stretching. They usually have a hook or eye form at each end for attachment.
- Compression coil springs, designed to resist being compressed. A typical use for compression coil springs is in car suspension systems.
- expansion coil spring.

Degradation

Many types of coil spring are wound in an annealed (soft) condition and then tempered to achieve their strength as a spring. Over time, this tempering can be lost and the spring will *sag* because it can no longer withstand the loads applied. Such springs can be *re-set* by annealing, returning to their original length (or deliberately setting them to a different length) and then re-tempering. Damage to springs, such as using oxy-acetylene to cut the end off a car suspension spring to lower a vehicle's ride height, can destroy the tempering in localised areas of the spring.

Springboard

A **springboard** or **diving board** is used for diving and is a board that is itself a spring, i.e. a linear flex-spring, of the cantilever type.

Springboards are commonly fixed by a hinge at one end (so they can be flipped up when not in use), and the other end usually hangs over a swimming pool, with a point midway between the hinge and the end resting on an adjustable fulcrum.

Springboard materials

Modern springboards are made out of a single-piece extrusion of aircraft-grade aluminum. The Maxiflex Model B, the board used in all major competitive diving events, is made out of such aluminum, and is heat treated for a yield strength of 50,000 psi. The slip-resistant surface of the board is created using an epoxy resin, finished with a laminate of flint silica and alumina in between the top coats of resin. This thermal-cured resin is aqua-colored to match the water of a clean pool.

Adjustment of the spring constant

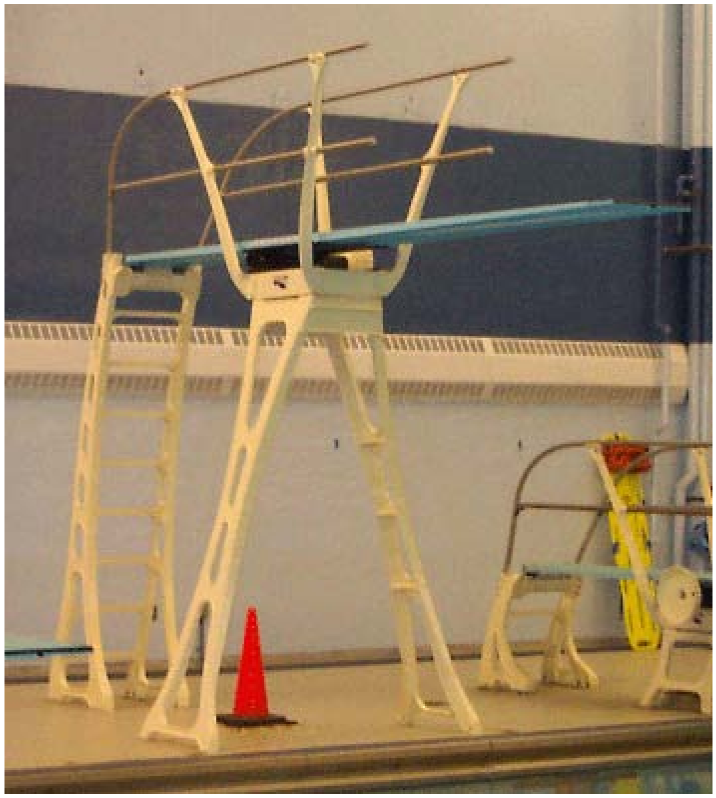
The spring constant of a springboard is usually adjusted by way of a fulcrum that is located approximately mid way along the springboard. Springboards are usually operated in a linear regime where they approximately obey Hooke's law. When loaded with a diver, the combination of the diver's approximately constant mass, and the constant stiffness of the spring(board) result in a resonance frequency that is adjustable by way of the spring constant (set by the fulcrum position). Since the resulting system is in an approximately linear regime, it may be modeled fairly accurately by a second order differential equation. Typically the resonance frequency can be adjusted over a range of a 2:1 or 3:1 ratio.

Counter-intuitive user-interface

The fulcrum usually travels over a range of approximately 0.75 metre (30 inches), and is set by way of a knob that is approximately 0.35 m (14 inches) in diameter. To stiffen the spring (as if tightening it), the knob is usually turned counter clockwise. This is counter intuitive, since usually things are tightened by turning clockwise. Additionally, if standing on the springboard, it is difficult to push the wheel with the foot, because the top of it needs to turn the other way from the way it moves. This is because the gearlike mechanism (usually a "soft gear" made of rubber) is on the board and not the base, so the wheel pivots against the board when rotated. Thus users often need to bend down and set the wheel, or come down from the board to set the wheel. Thus it would be much better if the gearing were on the base so that the wheel could be pushed with the foot, but tradition (consistency from board to board) dictactes maintaining a "backwards" convention., ,

- Note - Standing behind or in front of the knob, rather than directly above it, will give you better leverage to move the fulcrum. This is accomplished by holding on to the hand rails and leaning the body a few degrees, then placing your foot as low as possible on the knob. In this way, it is possible to move even the most difficult fulcrum.

Heights of springboards



Three and One Metre Springboards

Springboards are usually located either 1.0 metre or 3.0 metres above the water surface. It is very seldom that one is mounted at a height other than these two standard heights.

Historical heights of springboards

Some years ago, springboards, usually made of wood, were located at heights of either 10 feet (approximately 3m), or 20 feet (approx. 6m), above the water.

Home springboards

After an incident in Washington state in 1993, most US and other pool builders are reluctant to equip a residential swimming pool with a diving springboard so home diving pools are much less common these days. In the incident, 14-year-old Shawn Meneely made a "suicide dive" (his hands at his sides - so his head hit the bottom first) in a private swimming pool and was seriously injured (tetraplegic). Family lawyer Fred Zeder successfully sued the diving board manufacturer, the pool builder, and the National Spa and Pool Institute over the inappropriate depth of the pool. The NSPI had specified a minimum depth of 7 ft 6 in (2.55m) which proved to be insufficient in the above case. The pool into which Meneelly dived was not constructed to the published standards. The standards had changed after the diving board was installed on the non-compliant pool by the homeowner. But the courts held that the pool "was close enough" to the standards to hold NSPI liable. The multi-million dollar lawsuit was eventually settled in 2001 for $6,600,000USD ($US8,000,000 after interest was added) in favor of the plaintiff. The NSPI was held to be liable, and was financially strained by the case. It filed twice for Chapter 11 bankruptcy protection and was successfully reorganized into a new swimming pool industry association.

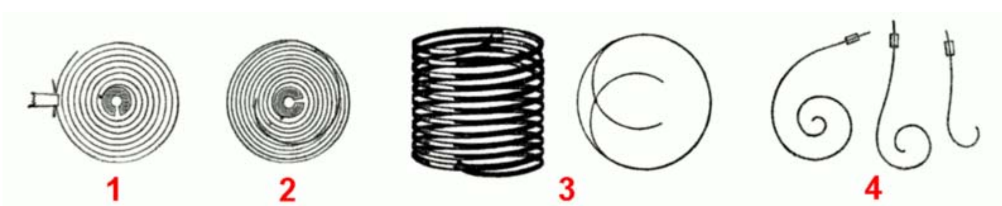
Chapter 3

Balance Spring

A **balance spring**, or **hairspring**, is a part used in mechanical timepieces. The balance spring, working together with the balance wheel, controls the speed of movement of the parts in the timepiece. The **regulator** lever is used to adjust the speed so that the timepiece keeps accurate time.

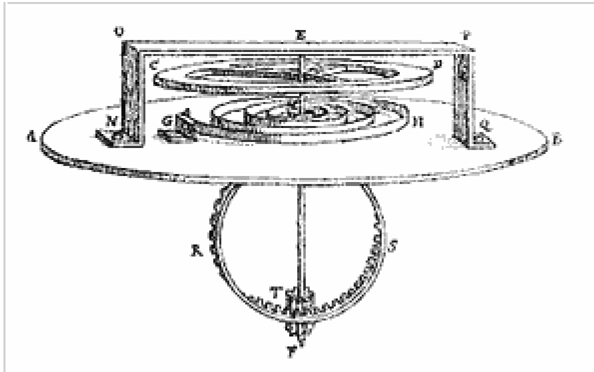
The invention of the balance spring in the 17th century greatly increased the accuracy of pocketwatches and other timepieces by controlling the speed of movement of the mechanism to a nearly-constant rate. Improvements since then made timepieces even more accurate by reducing the effect of temperature, and also the effect of the driving force, which for a mainspring decreases as the mainspring winds down.

The balance spring is a fine spiral or helical spring used in mechanical watches, marine chronometers, and other timekeeping mechanisms to control the rate of oscillation of the balance wheel. A balance spring usually has a regulator lever to adjust the rate of the timepiece, but other designs use adjusting screws on the balance wheel instead. The balance spring is an integral part of the balance wheel, because it reverses the direction of the balance wheel so it oscillates back and forth. The balance spring and balance wheel together form a harmonic oscillator, whose resonant period is resistant to changes from perturbing forces, which is what makes it a good timekeeping device.



Types of balance springs: (1) flat spiral, (2) Breguet overcoil, (3) chronometer helix, showing curving ends, (4) early balance springs.

History



Drawing of one of his first balance springs, attached to a balance wheel, by Christiaan Huygens, inventor of the balance spring, published in his letter in the *Journal des Sçavants* of 25 February 1675.

There is some dispute as to whether it was invented around 1660 by British physicist Robert Hooke or Dutch scientist Christian Huygens, with the likelihood being that Hooke first had the idea, but Huygens built the first functioning watch that used a balance spring. Before that time, balance wheels or foliots without springs were used in clocks and watches, but they were very sensitive to fluctuations in the driving force, causing the timepiece to slow down as the mainspring unwound. The introduction of the balance spring effected an enormous increase in the accuracy of pocketwatches, from perhaps several hours per day to 10 minutes per day, making them useful timekeepers for the first time. The first balance springs had only a few turns.

A few early watches had a Barrow regulator, which used a worm drive, but the first widely used regulator was invented by Thomas Tompion around 1680. In the Tompion regulator the curb pins were mounted on a semicircular toothed rack, which was adjusted by fitting a key to a cog and turning it. The modern regulator, a lever pivoted concentrically with the balance wheel, was patented by Joseph Bosley in 1755, but it didn't replace the Tompion regulator until the early 19th century.

Regulator

In order to adjust the rate, the balance spring usually has a *regulator*, a moveable lever with a narrow slit on the end through which the last turn of the spring passes. The portion of the spring after the slit is held stationary, so the slit controls the usable length of the spring. Moving the regulator slides the slit up or down the spring, changing its effective length. Moving it away from the spring's attachment point (stud) shortens the spring, making it stiffer, increasing the balance's oscillation rate, and making the timepiece gain time.

In older watches, the slit is the gap between two tiny pins, called the Curb Pins.

The regulator interferes slightly with the motion of the spring, causing inaccuracy, so precision timepieces like marine chronometers and some high end watches are *free sprung*, meaning they don't have a regulator. Instead, their rate is adjusted by timing screws on the balance wheel.

There are two principal types of Balance Spring Regulator.

- The Tompion Regulator, in which the Curb Pins are mounted on a sector-rack, moved by a pinion. The pinion is usually fitted with a graduated silver or steel disc.
- The Bosley Regulator, as described above, in which the Pins are mounted on a lever pivoted coaxially with the Balance, the extremity of the lever being able to be moved over a graduated scale. There are several variants of this, including the "Snail" regulator, in which the lever is sprung against a cam of spiral profile which can be turned, and the Micrometer, in which the lever is moved by a worm gear.

There is also a "Hog's Hair" or "Pig's Bristle" regulator, in which stiff fibres are positioned at the extremities of the Balance's arc, and bring it to a gentle halt before throwing it back. The Watch is accelerated by shortening the arc. This is not a Balance Spring Regulator, being used in the earliest Watches before the Balance Spring was invented.

There is also a Barrow Regulator, but this is really the earlier of the two principal methods of giving the Mainspring "set-up tension"; that required to keep the Fusee chain in tension but not enough to actually drive the Watch. Verge Watches can be regulated by adjusting the set-up tension, but if any of the previously described Regulators is present then this is not usually done.

Material

A number of materials have been used for balance springs. Early on, steel was used, but without any hardening or tempering process applied; as a result, these springs would gradually weaken and the watch would start losing time. Some watchmakers, for example John Arnold, used gold, which avoids the problem of corrosion, but retains the problem of gradual weakening. Hardened and tempered steel was first used by John Harrison and subsequently remained the material of choice until the 20th century.

In 1833, E. J. Dent (maker of the Great Clock of the Houses of Parliament) experimented with a glass Balance Spring. This was much less affected by heat than steel, reducing the Compensation required, and also didn't rust. Other trials with glass revealed that they were difficult and expensive to make, and there was a widespread opinion that they must be fragile. This latter objection is proved false by glass-fibre loft insulation and fibre-optic cables.

Effect of temperature

The modulus of elasticity of materials is dependent on temperature. For most materials, this temperature coefficient is large enough that variations in temperature significantly affect the timekeeping of a balance wheel and balance spring. The earliest makers of watches with balance springs, such as Robert Hooke and Christian Huygens observed this effect without finding a solution to it.

John Harrison, in the course of his development of the marine chronometer, solved the problem by a "compensation curb" -- essentially a bimetallic thermometer which adjusted the effective length of the balance spring as a function of temperature. While this scheme worked well enough to allow Harrison to meet the standards set by the Longitude Act, it was not widely adopted.

Around 1765, Pierre Le Roy (son of Julien Le Roy) invented the compensation balance, which became the standard approach for temperature compensation in watches and chronometers. In this approach, the shape of the balance is altered, or adjusting weights are moved on the spokes or rim of the balance, by a temperature-sensitive mechanism. This changes the moment of inertia of the balance wheel, and the change is adjusted such that it compensates for the change in modulus of elasticity of the balance spring. The compensating balance design of Thomas Earnshaw, which consists simply of a balance wheel with bimetallic rim, became the standard solution for temperature compensation.

Elinvar

While the compensating balance was effective as a way to compensate for the effect of temperature on the balance spring, it could not provide a complete solution. The basic design suffers from "middle temperature error": if the compensation is adjusted to be exact at extremes of temperature, then it will be slightly off at temperatures between those extremes. Various "auxiliary compensation" mechanisms were designed to avoid this, but they all suffer from being complex and hard to adjust.

Around 1900, a fundamentally different solution was created by Charles Édouard Guillaume, inventor of elinvar. This is a nickel-steel alloy with the property that the modulus of elasticity is essentially unaffected by temperature. A watch fitted with an elinvar balance spring requires either no temperature compensation at all, or very little. This simplifies the mechanism, and it also means that middle temperature error is eliminated as well, or at a minimum is drastically reduced.

Isochronism

A balance spring obeys Hooke's Law: the restoring torque is proportional to the angular displacement. When this property is exactly satisfied, the balance spring is said to be *isochronous*, and the period of oscillation is independent of the amplitude of oscillation. This is an essential property for accurate timekeeping, because no mechanical drive train can provide absolutely constant driving force. This is particularly true in watches and portable clocks which are powered by a mainspring, which provides a diminishing drive force as it unwinds. Another cause of varying driving force is friction, which varies as the lubricating oil ages.

Early watchmakers empirically found approaches to make their balance springs isochronous. For example, John Arnold in 1776 patented a helical (cylindrical) form of the balance spring, in which the ends of the spring were coiled inwards. In 1861 M. Phillips published a theoretical treatment of the problem. He demonstrated that a balance spring whose center of gravity coincides with the axis of the balance wheel is isochronous.

Period of oscillation

The balance spring is an essential part of the balance wheel; together they form a harmonic oscillator. The balance spring provides the linear restoring force that reverses the motion of the wheel so it oscillates back and forth. The motion of the balance wheel is approximately simple harmonic motion, i.e., a sinusoidal motion of constant period. Its resonant period is resistant to changes from perturbing forces, which is what makes it a good timekeeping device. The stiffness of the spring, its spring coefficient, κ in N-m/radian, along with the balance wheel's moment of inertia, I in kg-m^2, determines the wheel's oscillation period T in seconds:

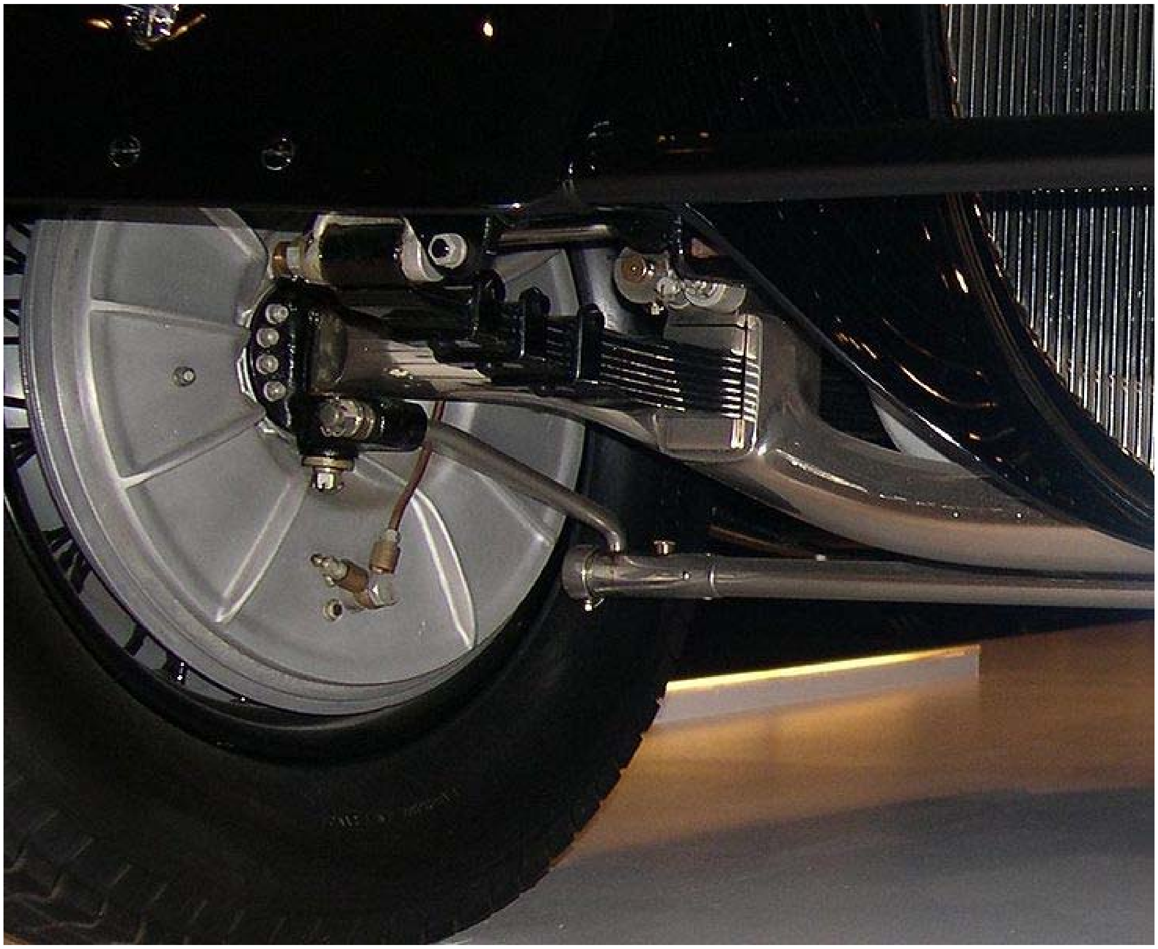
$$T = 2\pi\sqrt{I/\kappa}$$

This period controls the rate of the timepiece.

Chapter 4

Leaf Spring

A traditional semi-elliptical Hotchkiss leaf spring arrangement. On the left, the spring is connected to the frame through a shackle.



Quarter-elliptical spring in a 1937 Bugatti Type 57SC

Originally called *laminated* or *carriage spring*, a **leaf spring** is a simple form of spring, commonly used for the suspension in wheeled vehicles. It is also one of the oldest forms of springing, dating back to medieval times.

An advantage of a leaf spring over a helical spring is that the end of the leaf spring may be guided along a definite path.

Sometimes referred to as a **semi-elliptical spring** or **cart spring**, it takes the form of a slender arc-shaped length of spring steel of rectangular cross-section. The center of the arc provides location for the axle, while tie holes are provided at either end for attaching to the vehicle body. For very heavy vehicles, a leaf spring can be made from several leaves stacked on top of each other in several layers, often with progressively shorter leaves. Leaf springs can serve locating and to some extent damping as well as springing functions. While the interleaf friction provides a damping action, it is not well controlled and results in stiction in the motion of the suspension. For this reason manufacturers have experimented with mono-leaf springs.

A leaf spring can either be attached directly to the frame at both ends or attached directly at one end, usually the front, with the other end attached through a shackle, a short swinging arm. The shackle takes up the tendency of the leaf spring to elongate when compressed and thus makes for softer springiness. Some springs terminated in a concave end, called a *spoon end* (seldom used now), to carry a swivelling member.

History

There were a variety of leaf springs, usually employing the word "elliptical". "Elliptical" or "full elliptical" leaf springs referred to two circular arcs linked at their tips. This was joined to the frame at the top center of the upper arc, the bottom center was joined to the "live" suspension components, such as a solid front axle. Additional suspension components, such as trailing arms, would be needed for this design, but not for "semi-elliptical" leaf springs as used in the Hotchkiss drive. That employed the lower arc, hence its name. "Quarter-elliptic" springs often had the thickest part of the stack of leaves stuck into the rear end of the side pieces of a short ladder frame, with the free end attached to the differential, as in the Austin Seven of the 1920s. As an example of non-elliptic leaf springs, the Ford Model T had multiple leaf springs over its differential that were curved in the shape of a yoke. As a substitute for dampers (shock absorbers), some manufacturers laid non-metallic sheets in between the metal leaves, such as wood.

Leaf springs were very common on automobiles, right up to the 1970s in Europe and Japan and late 70's in America when the move to front wheel drive, and more sophisticated suspension designs saw automobile manufacturers use coil springs instead. Today leaf springs are still used in heavy commercial vehicles such as vans and trucks, SUVs, and railway carriages. For heavy vehicles, they have the advantage of spreading the load more widely over the vehicle's chassis, whereas coil springs transfer it to a single point. Unlike coil springs, leaf springs also locate the rear axle, eliminating the need for trailing arms and a Panhard rod, thereby saving cost and weight in a simple live axle rear suspension.

A more modern implementation is the parabolic leaf spring. This design is characterised by fewer leaves whose thickness varies from centre to ends following a parabolic curve. In this design, inter-leaf friction is unwanted, and therefore there is only contact between the springs at the ends and at the centre where the axle is connected. Spacers prevent contact at other points. Aside from a weight saving, the main advantage of parabolic springs is their greater flexibility, which translates into vehicle ride quality that approaches that of coil springs. There is a trade-off in the form of reduced load carrying capability, however. The characteristic of parabolic springs is better riding comfort and not as "stiff" as conventional "multi-leaf springs". It is widely used on buses for better comfort. A further development by the British GKN company and by Chevrolet with the Corvette amongst others, is the move to composite plastic leaf springs.

Typically when used in automobile suspension the leaf both supports an axle and locates/ partially locates the axle. This can lead to handling issues (such as 'axle tramp'), as the flexible nature of the spring makes precise control of the unsprung mass of the axle difficult. Some suspension designs which use leaf springs do not use the leaf to locate the axle and do not have this drawback. The Fiat 128's rear suspension is an example.

Manufacturing process

Multi-leaf springs are made as follows:

1. Shearing of flat bar
2. Center punching
3. End Heating process forming (hot & cold process)
    1. Eye Forming / Wrapper Forming
    2. Diamond cutting / end trimming / width cutting / end tapering
    3. End punching / end grooving / end bending / end forging / eye grinding
4. Heat treatment
    1. Center hole punching / nibbing
    2. Camber forming
    3. Quenching
    4. Tempering
5. Surface preparation
    1. Shot peening / stress peening
    2. Painting
6. Eye bush preparation process
    1. Eye reaming / eye boring
    2. Bush insertion
    3. Bush reaming
7. Assemble
    1. Presetting & load testing
    2. Paint touch-up
    3. Marking & packing

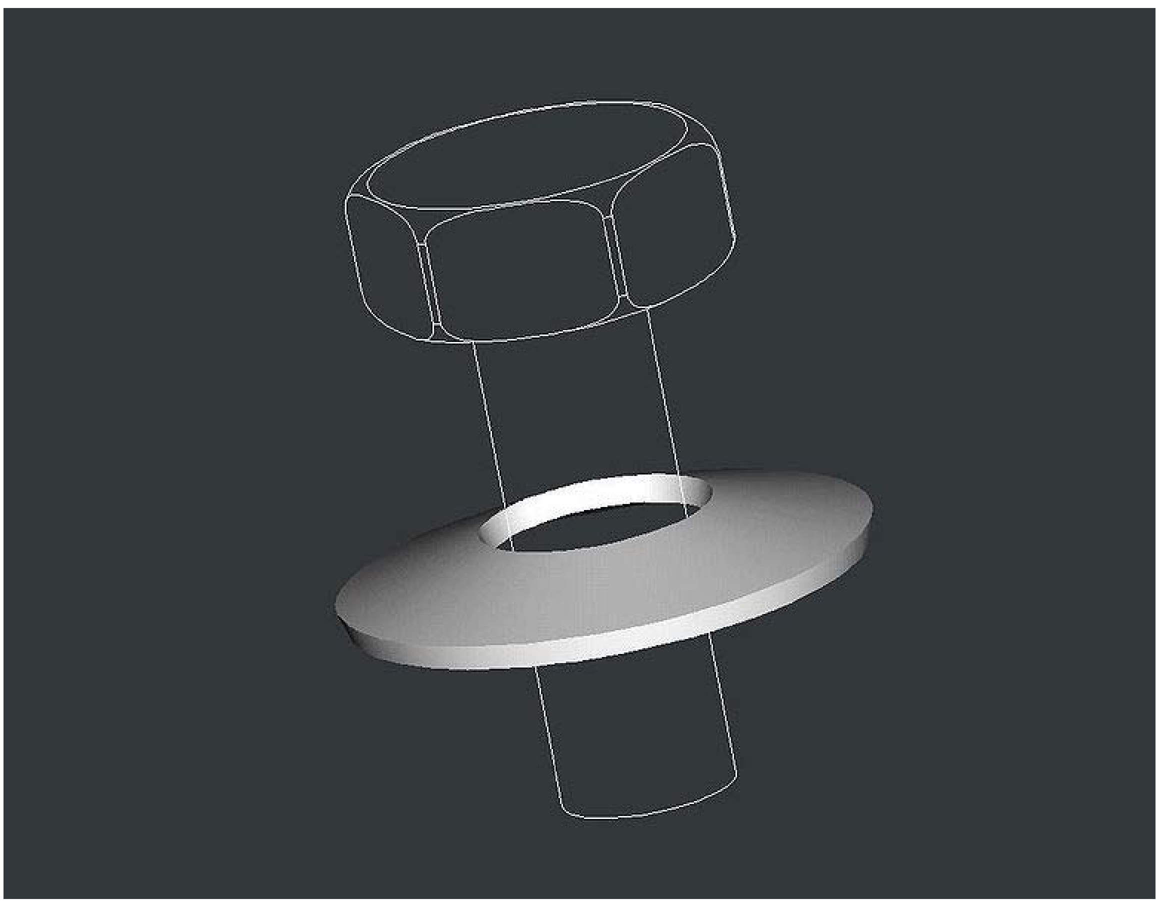
Use by Blacksmiths

Because leaf springs are made of relatively high quality steel, they are a favorite material for blacksmiths. In countries such as Nepal, Bangladesh and Pakistan where traditional blacksmiths still produce a large amount of the country's tools, leaf springs from scrapped cars are frequently used to make knives, kukris, and other tools. They are also commonly used by amateur and hobbyist blacksmiths in Western countries.

Chapter 5

Belleville Washer and Gas Spring

Belleville washer



Belleville washer

A **Belleville washer**, also known as a **coned-disc spring**, **conical spring washer**, **disc spring**, **Belleville spring** or **cupped spring washer**, is a type of spring shaped like a washer. It has a frusto-conical shape which gives the washer a spring characteristic. The Belleville name comes from the inventor Julian F. Belleville.

Design and use

Belleville washers are typically used as springs, or to apply a pre-load or flexible quality to a bolted joint or bearing.

Some properties of Belleville washers include: high fatigue life, better space utilization, low creep tendency, and high load capacity with a small spring deflection.

Belleville springs are also used in a number of landmines e.g. the American M15, M14, M1 and the Swedish Tret-Mi.59.

They may also be used as locking devices, but only in applications with low dynamic loads, such as down-tube shifters for bicycles. Belleville washers are seen on Formula One cars, as they provide extremely detailed tuning ability. The World War II-vintage German Junkers Ju 88 aircraft's single strut main gear made primary use of belleville washers as its main shock absorption mechanism. At least one modern aircraft design, the Cirrus SR2x series, uses a Belleville washer setup to damp out nose gear oscillations (or "shimmy").

Belleville washers have been used as return springs in artillery pieces, one example being the French Canet range of marine/coastal cannon from the late 1800's (75 mm, 120mm, 152 mm).

Another example where they aid locking is a joint that experiences a large amount of thermal expansion and contraction. They will supply the required pre-load, but the bolt may have an additional locking mechanism (like Loctite) that would fail without the Belleville.

Stacking

Multiple Belleville washers may be stacked to modify the spring constant or amount of deflection. Stacking in the same direction will add the spring constant in parallel, creating a stiffer joint (with the same deflection). Stacking in an alternating direction is the same as adding springs in series, resulting in a lower spring constant and greater deflection. Mixing and matching directions allow a specific spring constant and deflection capacity to be designed.

Example: 1 Spring is considered to be 1 in Parallel, 1 in Series. (This notation is needed for load calculations)

If n = # of springs in a stack, then: Parallel Stack (n in parallel, 1 in series) - Deflection is equal to that of one spring, Load is equal to that of n x 1 spring. i.e. Stack of 4 in parallel, 1 in series will have the same deflection as that of one spring and the load will be 4 times higher than that of one spring.

Series Stack (1 in parallel, n in series) - Deflection is equal to n x 1 spring, load is equal to that of one spring. i.e. Stack of 1 in parallel, 4 in series will have the same load of one spring and the deflection will be 4 times greater.

Performance considerations

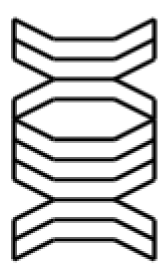
In a parallel stack, hysteresis (load losses) will occur due to friction between the springs. The hysteresis losses can be advantageous in some systems because of the added damping and dissipation of vibration energy. This loss due to friction can be calculated using hysteresis methods. Ideally, no more than 4 springs should be placed in parallel. If a greater load is required, then factor of safety must be increased in order to compensate for loss of load due to friction. Friction loss is not as much of an issue in series stacks

In a series stack, the deflection is not exactly proportional to the number of springs. This is because of a *bottoming out* effect when the springs are compressed to flat. The contact surface area increases once the spring is deflected beyond 95%. This decreases the moment arm and the spring will offer a greater spring resistance. Hysteresis can be used to calculate predicted deflections in a series stack. The number of springs used in a series stack is not as much of an issue as in parallel stacks.

Belleville washers are useful for adjustments because different thicknesses can be swapped in and out and they can be configured differently to achieve essentially infinite tunability of spring rate while only filling up a small part of the technician's tool box. They are ideal in situations where a heavy spring force is required with minimal free length and compression before reaching solid height. The downside, though, is weight, and they are severely travel limited compared to a conventional coil spring when free length is not an issue.

A similar device is a wave washer.

Calculation



2-3-1-2 stack of washers

If friction and bottoming-out effects are ignored, the spring rate of a stack of identical Belleville washers can be quickly approximated. Counting from one end of the stack, group by the number of adjacent washers in parallel. For example, in the stack of washers to the right, the grouping is 2-3-1-2, because there is a group of 2 washers in parallel, then a group of 3, then a single washer, then another group of 2.

The total spring coefficient is:

$$K = \frac{k}{\sum_{i=1}^{g} \frac{1}{n_i}}$$

$$K = \frac{k}{\frac{1}{2} + \frac{1}{3} + \frac{1}{1} + \frac{1}{2}}$$

$$K = \frac{3}{7}k$$

Where

- n_i = the number of washers in the ith group
- g = the number of groups
- k = the spring constant of one washer

So, a 2-3-1-2 stack (or, since addition is commutative, a 3-2-2-1 stack) gives a spring constant of 3/7 that of a single washer. These same 8 washers can be arranged in a 3-3-2 configuration (K = 6/7*k), a 4-4 configuration (K = 2*k), a 2-2-2-2 configuration (K = 1/2*k), and various other configurations. The number of unique ways to stack *n* washers is defined by the integer partition function *p*(*n*) and increases rapidly with large *n*, allowing fine-tuning of the spring constant. However, each configuration will have a different length, requiring the use of shims in most cases.

Gas spring

CGI of one type of gas spring

A **gas spring** is a type of spring that, unlike a typical metal spring, uses a compressed gas, contained in a cylinder and compressed by a piston, to exert a force. Gas springs are used in automobiles, where they are used to support the weight of doors while they are open. They are also used in furniture, medical, and aerospace applications. Gas springs are also used within the press tooling industry. These units are larger than the normal "strut" type units ranging in force from 2500N to 400,000N (Forty tonnes).

Theory

If a syringe plunger is squeezed with a closed nozzle, the resistance will rapidly rise. In a gas spring the volume is quite large compared to the diameter of the plunger and the gas, which may be dry air or nitrogen, is pre-compressed. Hence if a 1 square inch (6.45 square centimetre) plunger (radius 0.56 in or 1.43 cm) is used with a container with an internal pressure of 30 pounds per square inch (207 kPa), a thirty pound-force (130 N) spring will result. If the volume of gas is large, little increase in strength will result as the rod is pressed in. The gas volume can be altered to change this parameter. The standard gas equation is used to calculate the difference. Pressure x volume / temperature (must be kelvins or degrees Rankine) = pressure x volume / temperature. If the internal plunger has a diaphragm, which extends to the side of the gas tube it will not move. If a fine hole exists it will be a slow-dampened spring, for use on heavy doors and windows. If there is no diaphragm other than the washer to contain it, the result is a quick gas spring, as used on air rifles and recoil buffers. Reducing the gas volume and hence increasing its internal pressure by means of a movable end stop or allowing one tube to slide over another can achieve adjustability. The rod may be hollow by use of clever seals or can be multiple small-diameter rods. A small amount of oil is normally present. The gas may be introduced by Schrader-type valve, using a lip seal around the rod and forcing it to allow gas in by external over pressure or a shuttling O-ring system. Gas springs of high pressure contain a very large amount of energy and could be used as a power pack. In emergency use the gas may be introduced via a gas generator cell as used in air bags. Gas springs are used to operate the main valve on Formula 1 racing cars.

Other forms of standard gas springs

EasyStop adjustable retention force gas spring

A gas spring with adjustable retention force

- Push-in force can be adjusted
- Adjustment via bowden wire and turning knob
- Complete range is adjustable by a 270° turn

Facts:

```
Piston rod size:  10 mm
Cylinder size:    22 - 28 mm
Stroke:           10 - 700 mm
Force:            100-600 N
Retention force   0-1000 N (Friction — Push in direction)
Extended length:  2x stroke + 78 mm (10/22)
```

```
Extended length:  2x stroke + 89 mm (10/28)
```

Click and go gas spring is a form of lockable gas spring that allows a single touch to the release button and it will fully extend automatically.

- After short touch the gas spring pushes out over the whole stroke
- Lockable only in completely pushed-in condition
- All release systems usable

Facts:

```
Piston rod size: 8-10 mm
Cylinder size:   28 mm
Stroke:          10-300 mm (08/28),10-700 mm (10/28)
Force:           40-700 N (08/28),50-1300 N (10/28)
Extended length: 2x stroke + 78 mm (08/28)
Extended length: 2x stroke + 87 mm (10/28)
```

Aluminium gas spring is a standard gas spring with a weight reduction of about 52% when compared to a similar standard gas spring (with 150 mm stroke). Aluminium gas springs are highly corrosion resistant and have high flexural strength.

Facts:

```
Piston rod size:  8 mm
Cylinder size:    20 mm
Stroke:           10-300 mm
Force:            30-500 N
Extended length:  2x stroke + 49 mm
Progressivity:    Approx. 33%
Piston rod:       Aluminium hard anodized
Cylinder:         Aluminium powder coated
Connecting parts: Aluminium
```

Telescoping gas spring is a standard gas spring that offers double the stroke at a minimal compressed length. This is due to its ability to telescope out, similar to a telescoping antenna on an automobile. This gas spring is composed of one rod and two cylinders (The smaller of the two cylinders actually acts as a second rod extending in and out of the larger cylinder).

Facts:

Prototype example

```
Piston rod 1:        8 mm, CeramPro "coated" steel
Piston rod 2:        19 mm, CeramPro "coated" steel
Cylinder size:       40 mm, powder-coated steel
Compressed length:   480 mm
```

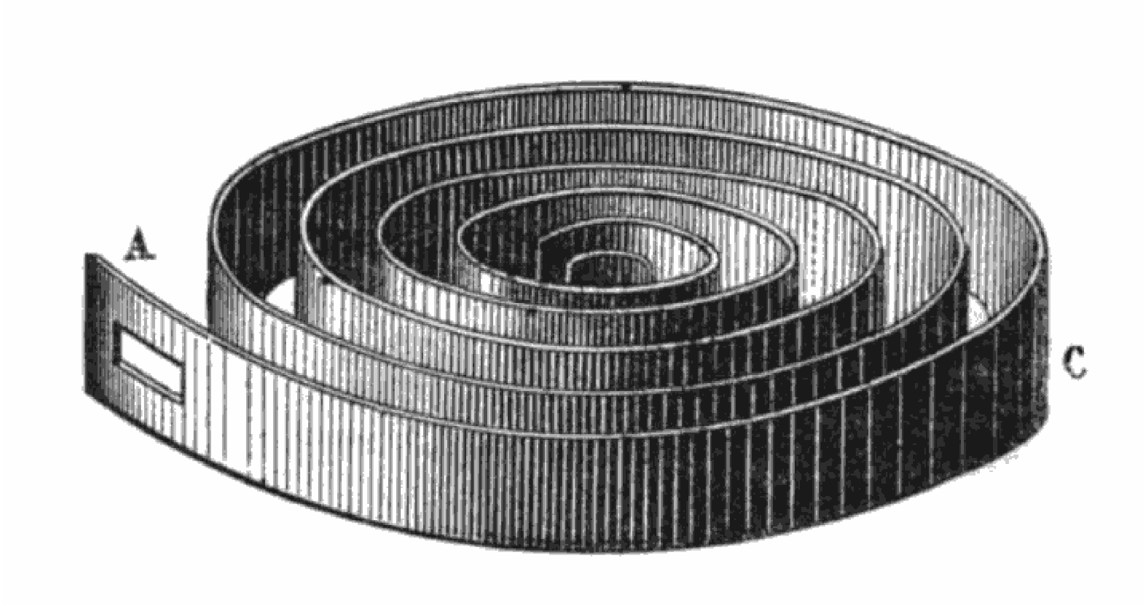
```
Stroke:              700 mm
Extended length:     1180 mm
Force:               100 N
Progressivity:       100%
```

Gas springs made by different manufacturers may use terminology or materials that differ from those noted here.

Chapter 6

Mainspring

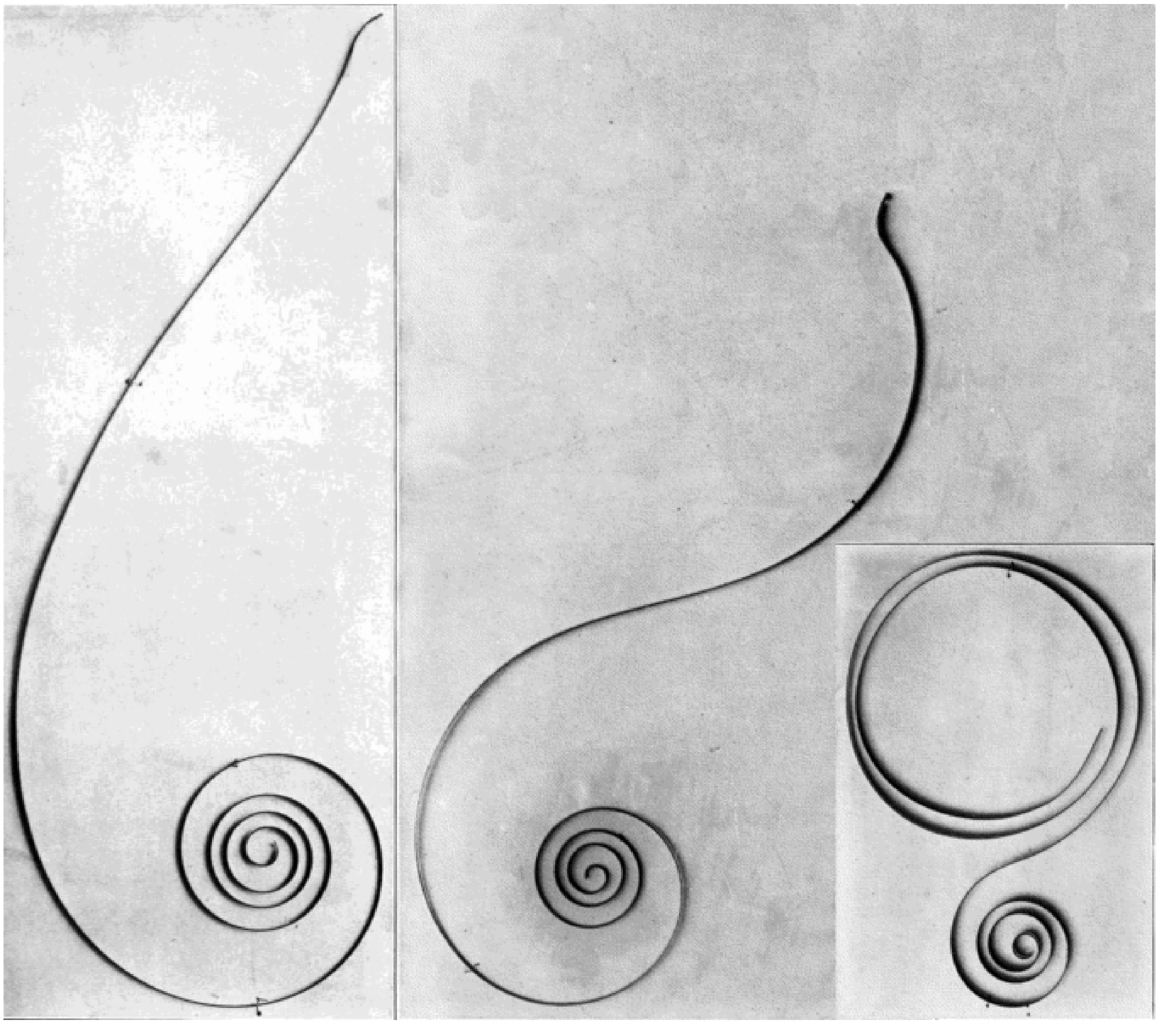
An uncoiled modern watch mainspring.



Clock mainspring

A **mainspring** is a spiral torsion spring of metal ribbon that is the power source in mechanical watches and some clocks. *Winding* the timepiece, by turning a knob or key, stores energy in the mainspring by twisting the spiral tighter. The force of the mainspring then turns the clock's wheels as it unwinds, until the next winding is needed. The adjectives **wind-up** and **spring-wound** refer to mechanisms powered by mainsprings, which also include kitchen timers, music boxes, wind-up toys and clockwork radios.

Modern mainsprings



Elgin pocketwatch mainsprings from around 1910, showing (l-r): spiral, semi-reverse, reverse.

A modern watch mainspring is a long strip of hardened and blued steel, or specialised steel alloy, 20-30 centimeters long and 0.05-0.2 millimeters thick. The mainspring in the common 1-day movement is calculated to enable the watch to run for 36 to 40 hours, i.e. with a power-reserve for 12 to 16 hours, which is the normal standard for hand-wound as well as self-winding watches. 8-Day movements provide power for at least 192 hours but use longer mainsprings and bigger barrels. Clock mainsprings are similar, only larger.

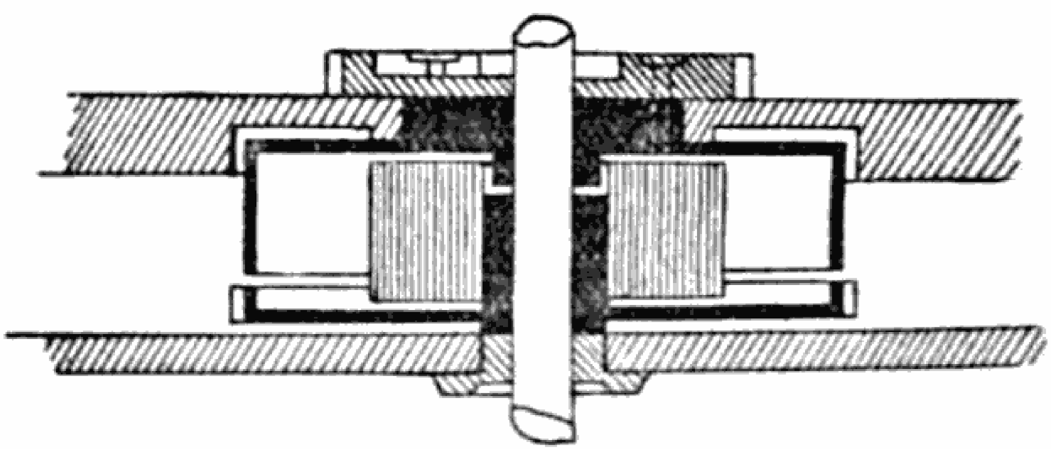
Since 1945, carbon steel alloys have been increasingly superseded by newer special alloys (iron, nickel and chromium with the addition of cobalt, molybdenum, or beryllium), and also by cold-rolled alloys ('structural hardening'). Known to watchmakers as 'white metal' springs (as opposed to blued carbon steel), these are stainless and have a higher elastic limit. They are less subject to permanent bending (becoming 'tired') and there is scarcely any risk of their breaking. Some of them are also practically non-magnetic.

In their relaxed form, mainsprings can have three distinct shapes:

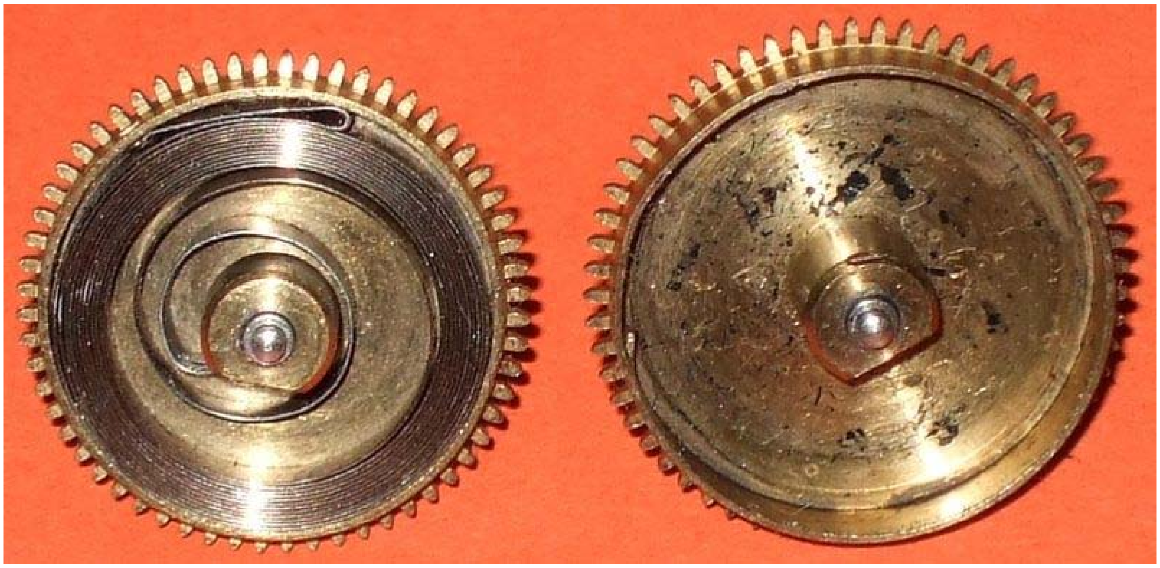
- Spiral coiled: i.e. coiled in the same direction throughout, viz. that of a spring inside the barrel
- Semi-reverse: The outer end of the spring is coiled in the reverse direction to form an angle less than 360 degrees.
- Reverse (resilient): the outer end of the spring is coiled in the reverse direction to form an angle exceeding 360 degrees.

The reverse coils provide extra force at the end of the running period, in order to keep the timepiece running at a constant rate to the end.

How they work



Cross section of a going barrel in a watch (mainspring fully wound).



Going barrel of a watch

The mainspring is coiled around an axle called the arbor, with the inner end hooked to it. In many clocks, the outer end is attached to a stationary post. The spring is wound up by turning the arbor, and after winding its force turns the arbor the other way to run the clock. The disadvantage of this arrangement is that while the mainspring is being wound, its drive force is removed from the clock movement, so the clock may stop. The winding mechanism must always have a ratchet attached, with a pawl (called by clockmakers the *click*) to prevent the spring from unwinding.

In the form used in modern watches, called the *going barrel*, the mainspring is coiled around an arbor and enclosed inside a cylindrical box called the barrel which is free to turn. The spring is attached to the arbor at its inner end, and to the barrel at its outer end. The attachments are small hooks or tabs, which the spring is hooked to by square holes in its ends, so it can be easily replaced.

The mainspring is wound by turning the arbor, but drives the watch movement by the barrel; this arrangement allows the spring to continue powering the watch while it is being wound. Winding the watch turns the arbor, which tightens the mainspring, wrapping it closer around the arbor. The arbor has a ratchet attached to it, with a click to prevent the spring from turning the arbor backward and unwinding. After winding, the arbor is stationary and the pull of the mainspring turns the barrel, which has a ring of gear teeth around it. This meshes with one of the clocks gears, usually the *center wheel* pinion and drives the wheel train. The barrel usually rotates once every 8 hours, so the common 40 hour spring requires 5 turns to unwind completely.

Hazards

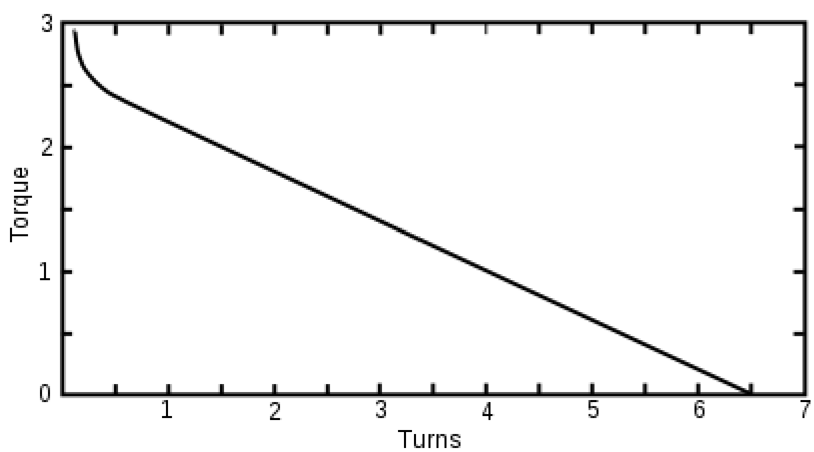
The mainspring contains a lot of energy. Clocks and watches have to be disassembled periodically for maintenance and repair, and if precautions are not taken the spring can release suddenly, causing serious injury. Mainsprings are 'let down' gently before servicing, by pulling the click back while holding the winding key, allowing the spring to slowly unwind. However, even in their 'let down' state, mainsprings in barrels contain dangerous residual tension. Watchmakers and clockmakers use a tool called a "mainspring winder" to safely install and remove them. Large mainsprings in clocks are immobilized by "mainspring clamps" before removal.

History

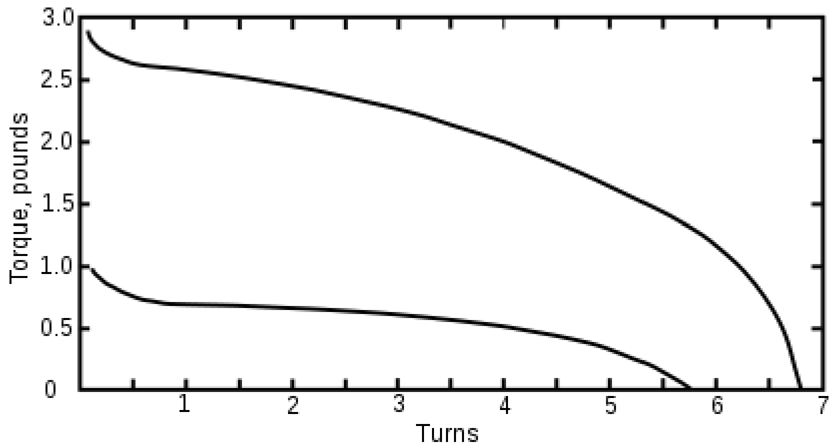
Mainsprings appeared in the first spring powered clocks, in 15th century Europe. Around 1400 coiled springs appeared in locks, and many early clockmakers were also locksmiths. Springs were applied to clocks to make them smaller and more portable than previous weight driven clocks, evolving into the first pocketwatches by 1600. Many sources erroneously credit the invention of the mainspring to the Nürnberg clockmaker Peter Henlein (or Henle, or Hele) around 1511. However, many descriptions from the 15th century of portable clocks 'without weights', and at least two surviving examples, show that spring driven clocks existed by the early years of that century. The oldest surviving clock powered by a mainspring is the *Burgunderuhr* (Burgundy Clock), an ornate, gilt spring driven chamber clock, currently at the Germanisches Nationalmuseum in Nurnberg, whose iconography suggests that it was made around 1430 for Philippe the Good, Duke of Burgundy..

The first mainsprings were made of steel without tempering or hardening processes. They didn't run very long, and had to be wound twice a day. Henlein was noted for making watches that would run 40 hours between windings.

Constant force from a spring



Torque curve of a mainspring. The force it provides decreases linearly as it unwinds.



Torque curves of mainsprings in going barrels (1879). The flatter central section provides more constant force during the running period, allowing the clock movement to keep better time.

A problem throughout the history of spring driven clocks and watches is that the force (torque) provided by a spring is not constant, but diminishes as the spring unwinds. Timepieces, however, have to run at a constant rate to keep accurate time. Timekeeping mechanisms are never isochronous; meaning their rate is affected by changes in the drive force. This was especially true of the primitive verge and foliot type used before the advent of the balance spring in 1657. So early clocks slowed down as the mainspring ran down.

Two solutions to this problem appeared in the early spring powered clocks in the 15th century: the *stackfreed* and the *fusee*. The stackfreed was an eccentric cam mounted on the mainspring arbor, with a spring-loaded roller that pressed against it. The cam was shaped so that early in the running period when the mainspring was pushing strongly, the stackfreed would provide an opposing force, while later when the mainspring was almost run down and pushing weakly, it would provide a helping force. The stackfreed added a lot of friction and probably reduced a clock's running time substantially; it was rarely used and was abandoned after about a century.

The fusee was a much longer lasting innovation. This was a cone-shaped pulley that was turned by a chain wrapped around the mainspring barrel. Its curving shape continuously changed the mechanical advantage of the linkage to even out the force of the mainspring as it ran down. Fusees became the standard method of getting constant torque from a mainspring. They were used in most spring driven clocks and watches from their first appearance until the 19th century when the going barrel took over, and in marine chronometers until the 1970s.

Another early device which helped even out the spring's force was *stopwork* or *winding stops*, which prevented the mainspring from being wound up all the way, and prevented it from unwinding all the way. The idea was to use only the central part of the spring's 'torque curve', where its force was more constant. The most common form was the Geneva stop or 'Maltese cross'. Stopwork isn't needed in modern watches.

A fourth device used in a few precision timepieces was the remontoire. This was a small secondary spring or weight which powered the timepiece's escapement, and was itself rewound periodically by the mainspring. This isolated the timekeeping element from the varying mainspring force.

16th century pocketwatch movement with stackfreed (near top).

The modern *going barrel*, invented in 1760 by Jean-Antoine Lépine, produces a constant force by simply using a longer mainspring than needed, and coiling it under tension in the barrel. In operation, only the inner turns of the spring are used. Mathematically, the tension creates a 'flat' section in the spring's 'torque curve' and only this flat section is used. In addition, the outer end of the spring is often given a 'reverse' curve, so it has an 'S' shape. This stores more tension in the spring's outer turns where it is available toward the end of the running period. The result is that the barrel provides approximately constant torque over the watch's designed running period; the torque doesn't decline until the mainspring has almost run down.

The built-in tension of the spring in the going barrel makes it hazardous to disassemble even when not wound up.

Broken mainsprings

Because they are subjected to constant stress cycles, up until the 1960s mainsprings generally broke from metal fatigue long before other parts of the timepiece. They were considered expendable items. This often happened at the end of the winding process, when the spring is wound as tightly as possible around the arbor, with no space between the coils. When manually winding, it is easy to reach this point unexpectedly and put excessive pressure on the spring. Broken mainsprings were the largest cause of watch repairs until the 1960s. Since then, the improvements in spring metallurgy mentioned above have made broken mainsprings rare.

Even if the spring didn't break, too much force caused another problem. Since no more slack was left in the spring, the pressure of the last turn of the winding knob put the spring under excessive tension, which was locked in by the last click of the ratchet. So the watch ran with excessive drive force for several hours, until the extra tension in the end of the spring was relieved. This caused the balance wheel to rotate too far and 'knock', and the watch to gain time. In older watches this was prevented with 'stopwork'. In modern watches this is prevented by designing the 'click' with some 'recoil' (backlash), to allow the arbor to rotate backward after winding by about two ratchet teeth, enough to remove excess tension.

Motor or safety barrel

Around 1900, when broken watchsprings were more of a problem, some pocketwatches used a variation of the going barrel called the *motor barrel* or *safety barrel*. Mainsprings usually broke at their attachment to the arbor, where bending stresses are greatest. When the mainspring broke, the outer part recoiled and the momentum spun the barrel in the reverse direction. This applied great force to the delicate wheel train and escapement, often breaking pivots and jewels.

In the motor barrel, the functions of the arbor and barrel were reversed from the going barrel. The mainspring was wound by the barrel, and turned the arbor to drive the wheel train. Thus if the mainspring broke, the destructive recoil of the barrel would be applied not to the wheel train but to the winding mechanism, which was robust enough to take it.

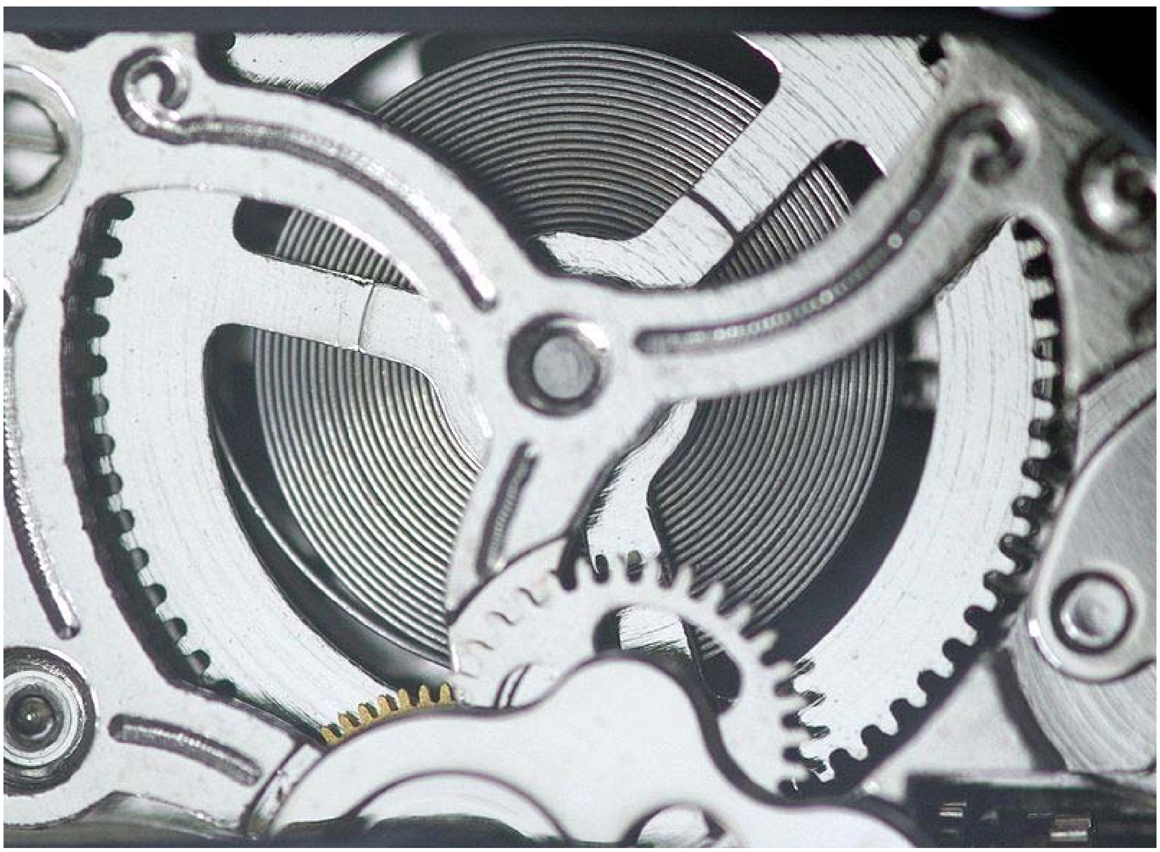
Safety pinion

A *safety pinion* was an alternate means of protection, used with the going barrel. In this, the center wheel pinion, which the barrel gear engages, was attached to its shaft with a reverse screw thread. If the spring broke, the reverse recoil of the barrel, instead of being passed on to the gear train, would simply unscrew the pinion.

The myth of 'overwinding'

Watches are often found stopped with the mainspring fully wound, which led to a myth that winding a watch all the way up damages it. What actually happens is that as time passes and the watch movement collects dirt and the oil dries up, friction increases, so that the mainspring doesn't have the force to turn the watch until the end of its running period. If the owner continues to wind and use the watch, eventually the friction force reaches the 'flat' part of the torque curve, and quickly a point is reached where the mainspring doesn't have the force to run the watch even at full wind, so the watch stops with the mainspring fully wound. The watch needs service, but the problem is caused by a dirty movement or other defect, not 'overwinding'.

Self-winding watches and 'unbreakable' mainsprings



The mainspring of an automatic watch. The spring isn't firmly mounted on the left side, and will slip when fully wound.

Self-winding or automatic watches, introduced widely in the 1950s, use the natural motions of the wrist to keep the mainspring wound. A semicircular weight, pivoted at the center of the watch, rotates with each wrist motion. A winder mechanism uses rotations in both directions to wind the mainspring.

In automatic watches, motion of the wrist could continue winding the mainspring until it broke. This is prevented with a slipping clutch device. The outer end of the mainspring, instead of attaching to the barrel, is attached to a circular expansion spring called the *bridle* that presses against the inner wall of the barrel, which has serrations or notches to hold it. During normal winding the bridle holds by friction to the barrel, allowing the mainspring to wind. When the mainspring reaches its full tension, its pull is stronger than the bridle. Further rotation of the arbor causes the bridle to slip along the barrel, preventing further winding. In watch company terminology, this is often misleadingly referred to as an 'unbreakable mainspring'.

'Tired' or 'set' mainsprings

After decades of use, mainsprings in older timepieces are found to deform slightly and lose some of their force, becoming 'tired' or 'set'. This condition is mostly found in springs in barrels. It causes the running time between windings to decrease. During servicing the mainspring should be checked for 'tiredness' and replaced if necessary. The British Horological Institute suggests these tests:

- In a mainspring barrel, when unwound and relaxed, most of a healthy spring's turns should be pressed flat against the wall of the barrel, with only 1 or 2 turns spiralling across the central space to attach to the arbor. If more than 2 turns are loose in the center, the spring may be 'tired'; with 4 or 5 turns it definitely is 'tired'.
- When removed from the barrel, if the diameter of the relaxed spring lying on a flat surface is less than 2½ times the barrel diameter, it is 'tired'.

Power reserve indicator

The power reserve is at the 6 position on this automatic watch. Here it is indicating that 25 out of 40 hours remain

Some high grade watches have an extra dial on the face indicating how much power is left in the mainspring, often graduated in hours the watch has left to run. Since both the arbor and the barrel turn, this mechanism requires a differential gear that measures how far the arbor has been turned, compared to the barrel.

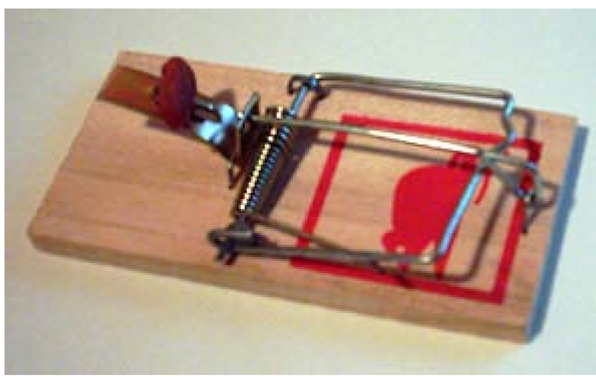
Unusual forms of mainspring

A mainspring is usually a coiled metal spring, however there are exceptions:

- The wagon spring clock: During a brief time in American clockmaking history coilable spring steel was not available in the USA and inventive clockmakers built clocks powered by a stack of leafsprings, similar what has traditionally served as a suspension spring for wagons.
- Other spring types are conceivable and have been used occasionally on experimental timepieces, such as e.g. torsion springs.
- Occasionally one finds an odd clock with a spring made of material other than metal, such as e.g. synthetic elastic materials.

Chapter 7

Torsion Spring



A mousetrap powered by a helical torsion spring

A **torsion spring** is a spring that works by torsion or twisting; that is, a flexible elastic object that stores mechanical energy when it is twisted. The amount of force (actually torque) it exerts is proportional to the amount it is twisted. There are two types. A **torsion bar** is a straight bar of metal or rubber that is subjected to twisting (shear stress) about its axis by torque applied at its ends. A more delicate form used in sensitive instruments, called a **torsion fiber** consists of a fiber of silk, glass, or quartz under tension, that is twisted about its axis. The other type, a **helical torsion spring**, is a metal rod or wire in the shape of a helix (coil) that is subjected to twisting about the axis of the coil by sideways forces (bending moments) applied to its ends, twisting the coil tighter. This terminology can be confusing because in a helical torsion spring the forces acting on the wire are actually bending stresses, not torsional (shear) stresses.

Torsion coefficient

As long as they are not twisted beyond their elastic limit, torsion springs obey an angular form of Hooke's law:

$$\tau = -\kappa\theta$$

where τ is the torque exerted by the spring in newton-meters, and θ is the angle of twist from its equilibrium position in radians. κ is a constant with units of newton-meters / radian, variously called the spring's **torsion coefficient**, **torsion elastic modulus**, **rate**, or just **spring constant**, equal to the torque required to twist the spring through an angle of 1 radian. It is analogous to the spring constant of a linear spring.

The energy *U*, in joules, stored in a torsion spring is:

$$U = \frac{1}{2}\kappa\theta^2$$

Uses

Some familiar examples of uses are the strong helical torsion springs that operate clothespins and traditional springloaded-bar type mousetraps. Other uses are in the large coiled torsion springs used to counter-balance the weight of garage doors, and a similar system is used to assist in opening the trunk (boot) cover on some sedans. Small coiled torsion springs are often used to operate pop-up doors found on small consumer goods like digital cameras and compact disc players. Other more specific uses:

- *Torsion bar suspension* has been used in vehicle suspension at least from the time of the 1934 Citroen Traction Avant. The Volkswagen Beetle and Porsche 911 are probably the most famous examples, but it has also been the dominant heavy armored vehicle suspension since World War II, and is now common in SUVs.

- The *sway bar* used in many vehicle suspension systems also uses the torsion spring principle.

- The *torsion pendulum* used in torsion pendulum clocks is a wheel-shaped weight suspended from its center by a wire torsion spring. The weight rotates about the axis of the spring, twisting it, instead of swinging like an ordinary pendulum. The force of the spring reverses the direction of rotation, so the wheel oscillates back and forth, driven at the top by the clock's gears.

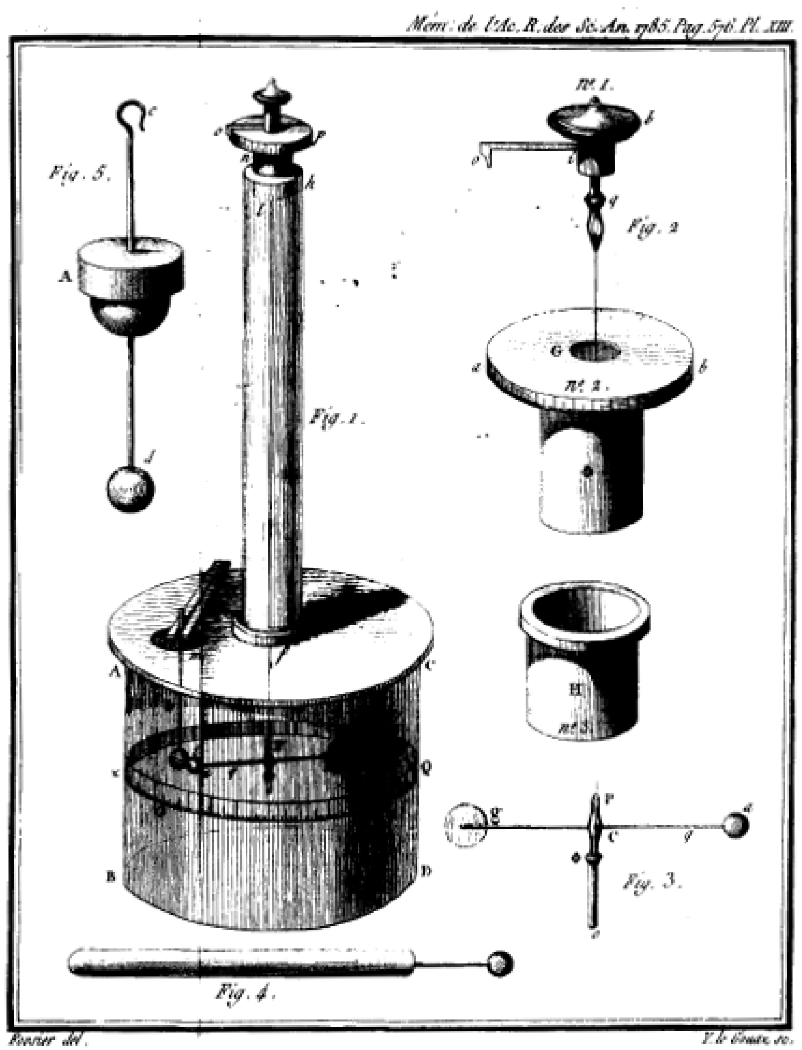
- The *torsion catapult* or mangonel is a medieval siege engine invented by the ancient Greeks. It uses a torsion spring consisting of twisted ropes to swing an arm that throws a heavy missile at the enemy with great force.

- The *balance spring* or hairspring in mechanical watches is a fine spiral-shaped torsion spring that pushes the balance wheel back toward its center position as it rotates back and forth. The balance wheel and spring function similarly to the torsion pendulum above in keeping time for the watch.

- The *D'Arsonval movement* used in mechanical pointer-type meters to measure electrical current is a type of torsion balance. A coil of wire attached to the pointer twists in a magnetic field against the resistance of a torsion spring. Hooke's law ensures that the angle of the pointer is proportional to the current.

- A *DMD* or digital micromirror device chip is at the heart of many video projectors. It uses hundreds of thousands of tiny mirrors on tiny torsion springs fabricated on a silicon surface to reflect light onto the screen, forming the image.

Torsion balance



Drawing of Coulomb's torsion balance. From Plate 13 of his 1785 memoir.

The **torsion balance**, also called **torsion pendulum**, is a scientific apparatus for measuring very weak forces, usually credited to Charles-Augustin de Coulomb, who invented it in 1777, but independently invented by John Michell sometime before 1783. Its most well-known uses were by Coulomb to measure the electrostatic force between charges to establish Coulomb's Law, and by Henry Cavendish in 1798 in the Cavendish experiment to measure the gravitational force between two masses to calculate the density of the Earth, leading later to a value for the gravitational constant.

The torsion balance consists of a bar suspended from its middle by a thin fiber. The fiber acts as a very weak torsion spring. If an unknown force is applied at right angles to the ends of the bar, the bar will rotate, twisting the fiber, until it reaches an equilibrium where the twisting force or torque of the fiber balances the applied force. Then the magnitude of the force is proportional to the angle of the bar. The sensitivity of the instrument comes from the weak spring constant of the fiber, so a very weak force causes a large rotation of the bar.

In Coulomb's experiment, the torsion balance was an insulating rod with a metal-coated ball attached to one end, suspended by a silk thread. The ball was charged with a known charge of static electricity, and a second charged ball of the same polarity was brought near it. The two charged balls repelled one another, twisting the fiber through a certain angle, which could be read from a scale on the instrument. By knowing how much force it took to twist the fiber through a given angle, Coulomb was able to calculate the force between the balls. Determining the force for different charges and different separations between the balls, he showed that it followed Coulomb's law.

To measure the unknown force, the spring constant of the torsion fiber must first be known. This is difficult to measure directly because of the smallness of the force. Cavendish accomplished this by a method widely used since: measuring the resonant vibration period of the balance. If the free balance is twisted and released, it will oscillate slowly clockwise and counterclockwise as a harmonic oscillator, at a frequency that depends on the moment of inertia of the beam and the elasticity of the fiber. Since the inertia of the beam can be found from its mass, the spring constant can be calculated.

Coulomb first developed the theory of torsion fibers and the torsion balance in his 1785 memoir, *Recherches theoriques et experimentales sur la force de torsion et sur l'elasticite des fils de metal &c*. This led to its use in other scientific instruments, such as galvanometers, and the Nichols radiometer which measured the radiation pressure of light. In the early 1900s gravitational torsion balances were used in petroleum prospecting. Today torsion balances are still used in physics experiments. In 1987, gravity researcher A.H. Cook wrote:

The most important advance in experiments on gravitation and other delicate measurements was the introduction of the torsion balance by Michell and its use by Cavendish. It has been the basis of all the most significant experiments on gravitation ever since.

Torsional harmonic oscillators

For definition of terms see end of section

Torsion balances, torsion pendulums and balance wheels are examples of torsional harmonic oscillators that can oscillate with a rotational motion about the axis of the torsion spring, clockwise and counterclockwise, in harmonic motion. Their behavior is analogous to translational spring-mass oscillators. The general equation of motion is:

$$I\frac{d^2\theta}{dt^2} + C\frac{d\theta}{dt} + \kappa\theta = \tau(t)$$

If the damping is small, $C \ll \sqrt{\kappa I}$, as is the case with torsion pendulums and balance wheels, the frequency of vibration is very near the natural resonance frequency of the system:

$$f_n = \frac{\omega_n}{2\pi} = \frac{1}{2\pi}\sqrt{\kappa/I}$$

The general solution in the case of no drive force ($\tau = 0$), called the transient solution, is:

$$\theta = Ae^{-\alpha t}\cos(\omega t + \phi)$$

where:

$$\alpha = C/2I$$
$$\omega = \sqrt{\omega_n^2 - \alpha^2} = \sqrt{\kappa/I - (C/2I)^2}$$

Applications

The balance wheel of a mechanical watch is a harmonic oscillator whose resonance frequency f_n sets the rate of the watch. The resonance frequency is regulated, first coarsely by adjusting I with weight screws set radially into the rim of the wheel, and then more finely by adjusting κ with a regulating lever that changes the length of the balance spring.

In a torsion balance the drive torque is constant and equal to the unknown force to be measured F, times the moment arm of the balance beam L, so $\tau(t) = FL$. When the oscillatory motion of the balance dies out, the deflection will be proportional to the force:

$$\theta = FL/\kappa$$

To determine F it is necessary to find the torsion spring constant κ. If the damping is low, this can be obtained by measuring the natural resonance frequency of the balance, since the moment of inertia of the balance can usually be calculated from its geometry, so:

$$\kappa = (2\pi f_n)^2 I$$

In measuring instruments, such as the D'Arsonval ammeter movement, it is often desired that the oscillatory motion die out quickly so the steady state result can be read off. This is accomplished by adding damping to the system, often by attaching a vane that rotates in a fluid such as air or water (this is why magnetic compasses are filled with fluid). The value of damping that causes the oscillatory motion to settle quickest is called the critical damping C_c:

$$C_c = 2\sqrt{\kappa I}$$

Definition of terms

Term	Unit	Definition
θ	radians	Angle of deflection from rest position
I	$\mathrm{kg\,m^2}$	Moment of inertia
C	$\mathrm{kg\,m^2\,s^{-1}\,rad^{-1}}$	Rotational friction (damping)
κ	$\mathrm{N\,m\,rad^{-1}}$	Coefficient of torsion spring
τ	$\mathrm{N\,m}$	Drive torque
f_n	Hz	Undamped (or natural) resonance frequency
ω_n	$\mathrm{rad\,s^{-1}}$	Undamped resonance frequency in radians
f	Hz	Damped resonance frequency
ω	$\mathrm{rad\,s^{-1}}$	Damped resonance frequency in radians
α	$\mathrm{s^{-1}}$	Reciprocal of damping time constant
ϕ	rad	Phase angle of oscillation
L	m	Distance from axis to where force is applied

Chapter 8

Corvette Leaf Spring

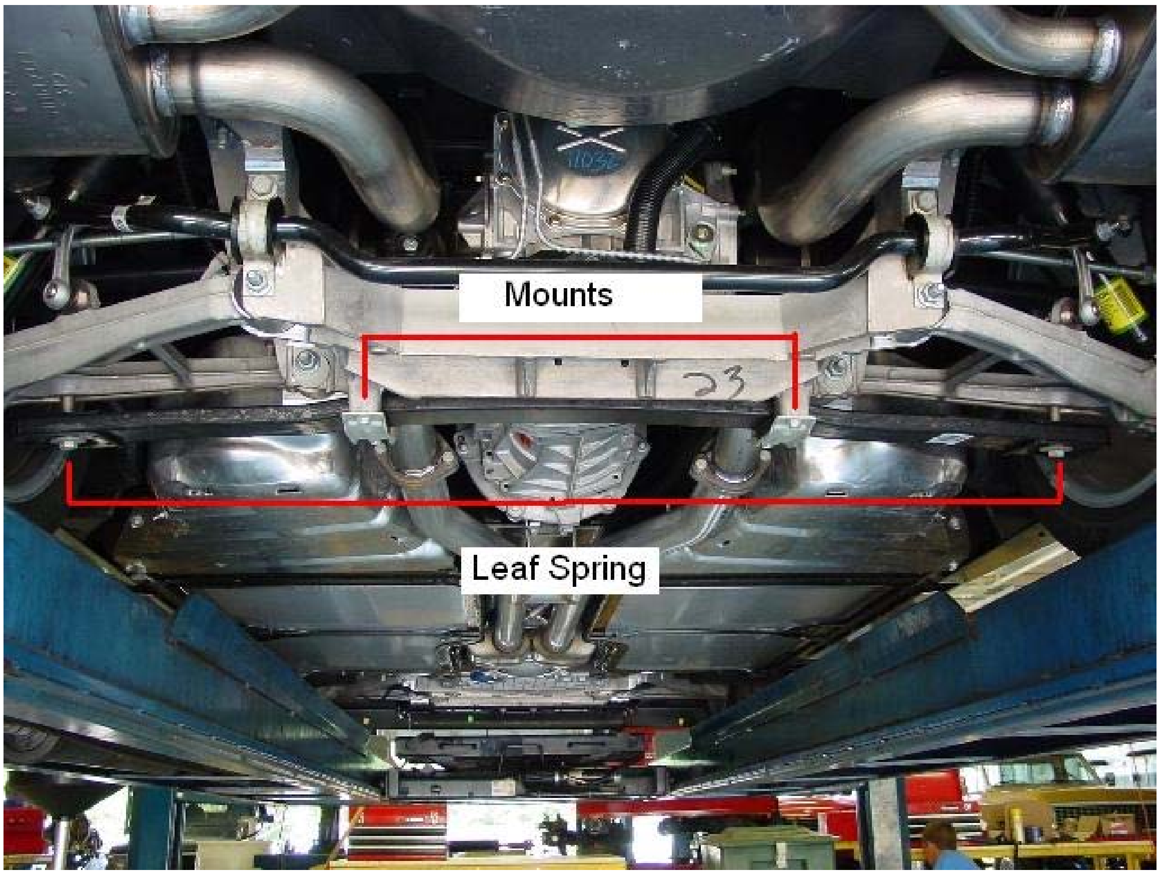
Since 1963, transverse **leaf springs** have been an integral part of the suspension of **GM'**s **Chevrolet Corvette**.

Traditional use of leaf springs

A traditional leaf spring arrangement.

A leaf spring is a long, flat, thin, and flexible piece of spring steel or composite material that resists bending. The basic principles of leaf spring design and assembly are relatively simple, and leafs have been used in various capacities since medieval times. Most heavy duty vehicles today use two sets of leaf springs per solid axle, mounted perpendicularly to support the weight of the vehicle. This Hotchkiss system requires that each leaf set act as both a spring and a horizontally stable link. Because leaf sets lack rigidity, such a dual-role is only suited for applications where load-bearing capability is more important than precision in suspension response.

Leaf springs on the Corvette



The C5 Corvette's rear suspension.

All six generations of the Corvette have used leaf springs in some capacity. The basic arrangement for each generation is listed as follows:

- C1 (1953–1962):

  Front: Independent unequal-length double wishbones with coil springs.
  Rear: Rigid axle supported by leaf springs and longitudinal control links.

- C2 (1963–1967), C3 (1968–1982):

Front: Independent unequal-length double wishbones with coil springs.
Rear: Independent suspension with trailing and lateral links supported by a centrally mounted leaf spring.

- C4 (1984–1996):

  Front: Independent unequal-length double wishbones with transverse fiberglass mono-leaf spring mounted to allow for anti-roll effect.
  Rear: Independent suspension with trailing and lateral links supported by a centrally mounted fiberglass mono-leaf spring.

- C5 (1997–2004), C6 (2005–):

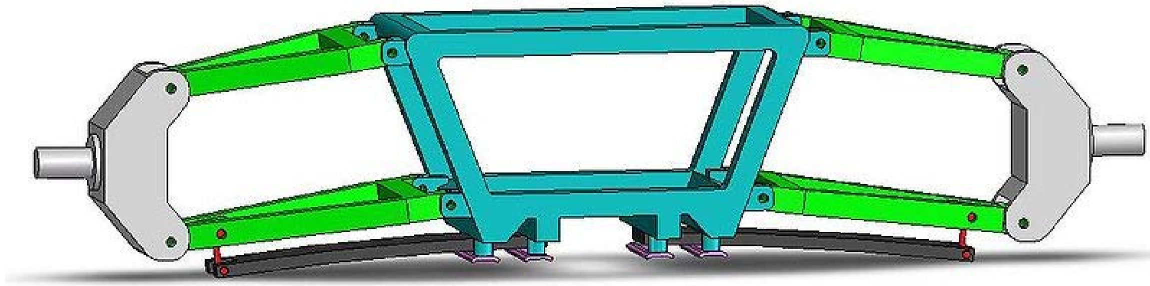
  Front: Independent unequal-length double wishbones with transverse fiberglass mono-leaf spring mounted to allow for anti-roll effect.
  Rear: Independent unequal length double wishbones with transverse fiberglass mono-leaf spring mounted to allow for anti-roll effect.

In the C2 and subsequent generations, a leaf spring is mounted transversely in the chassis and used in conjunction with several independent suspension designs. Common to these post-C1 Corvettes, the leaf acts only as a spring, and not a suspension arm or a link. Because it is not required to stabilize the wheels, the leaf functions in much the same manner as a coil spring. This configuration obviates the drawbacks and imprecision associated with leaf springs in a traditional Hotchkiss suspension layout.

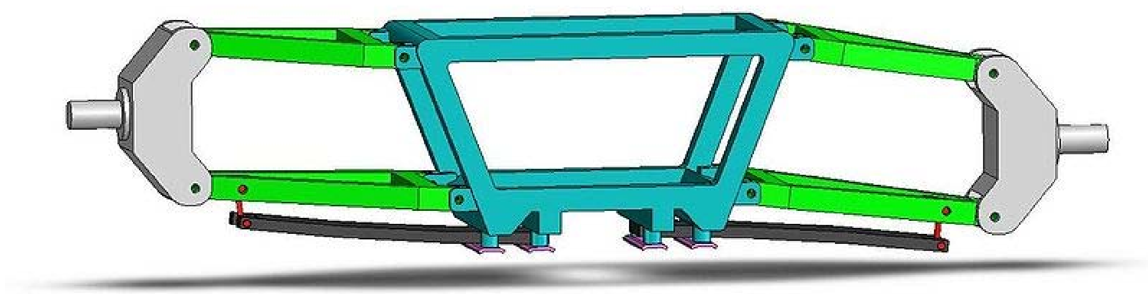
Motion of a transverse leaf spring

The following images show the movements of an independent suspension using a transverse leaf spring. For all images:

- The suspension arms are green
- The chassis is blue
- The uprights are gray
- Leaf springs are dark gray
- Pivot links connecting the ends of the springs to the suspension arms are red

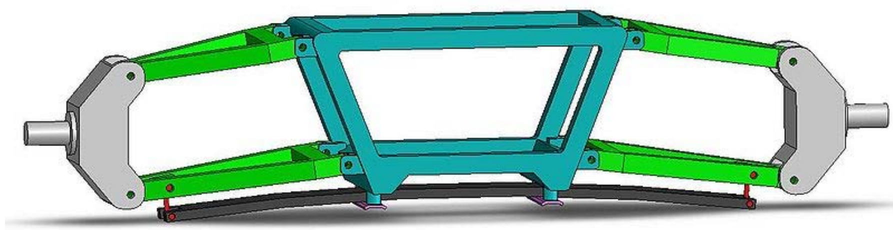


1 - A transverse leaf spring suspension at rest, with separate right and left springs.

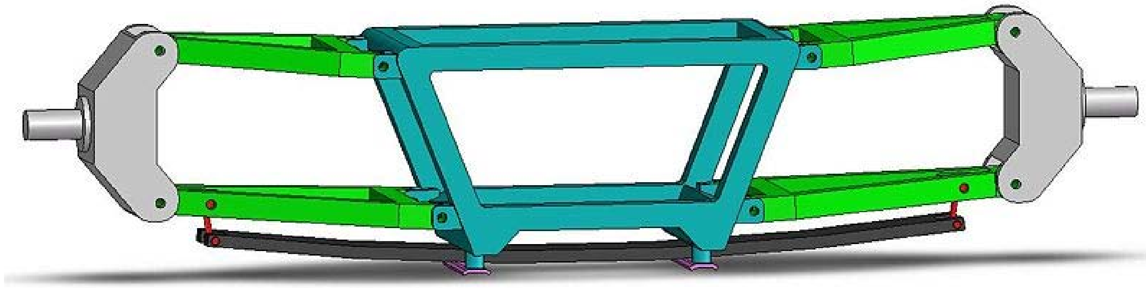


2 - The same split-spring configuration with the left wheel in compression.

Illustrations #1 and #2 show independent left and right leaf springs mounted rigidly to a chassis. In the first illustration, the suspension is at rest. As a left wheel moves up in the second illustration, the left spring flexes upward, but the right spring remains unaffected. Because the two springs are not connected, the movement of one wheel has no effect on the spring rate of the opposite wheel. While the C2, C3, and C4 Corvettes used a continuous spring instead of the split spring of the illustration, left and right spring rates remained independent because the spring was rigidly mounted at its center to the chassis.

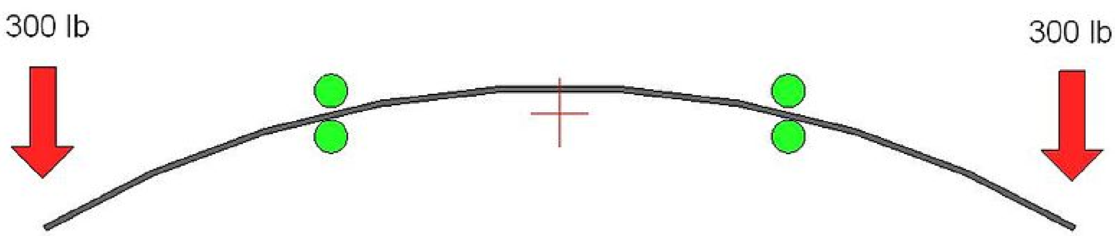


3 - A single transverse leaf spring suspension similar to that used on the C5 and C6 Corvette.

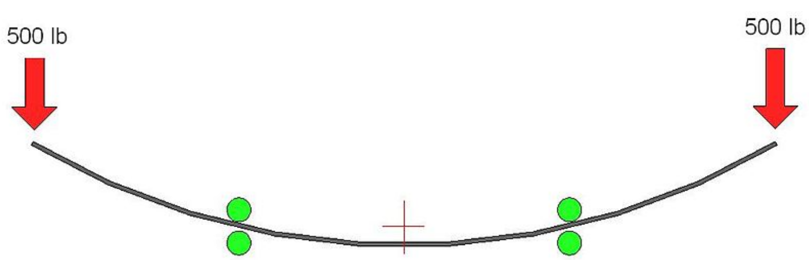


4 - The same single-leaf suspension with both wheels compressed upward.

Illustrations #3 and #4 show an independent suspension with a single transverse leaf spring, an arrangement similar to that used on the C5 and C6 Corvettes and the front of the C4 Corvette. While at rest in illustration #3, the leaf forms a symmetric arc between the left and right sides of the suspension with equal force applied to each. Under the compression of both wheels in illustration #4, the widely-spaced chassis mounts allow the spring to pivot; the ends of the spring flex upward and the center moves down. Spring force remains even between both sides.



At static ride height the leaf spring applies the same 300lb to each side of the suspension.

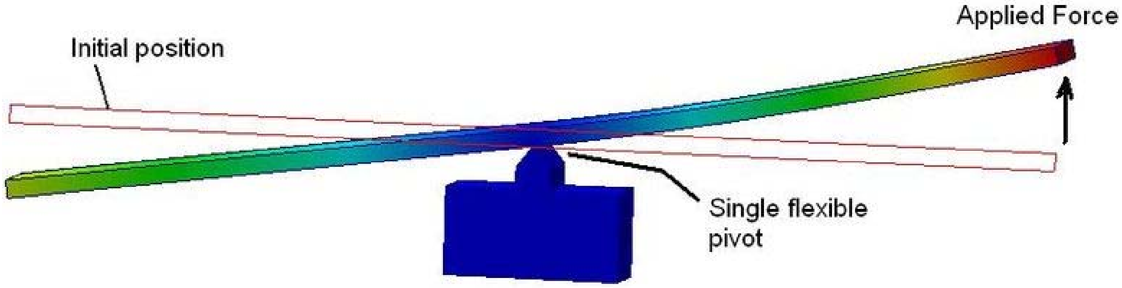


In compression the spring force has increased to 500lb but is still even between both sides

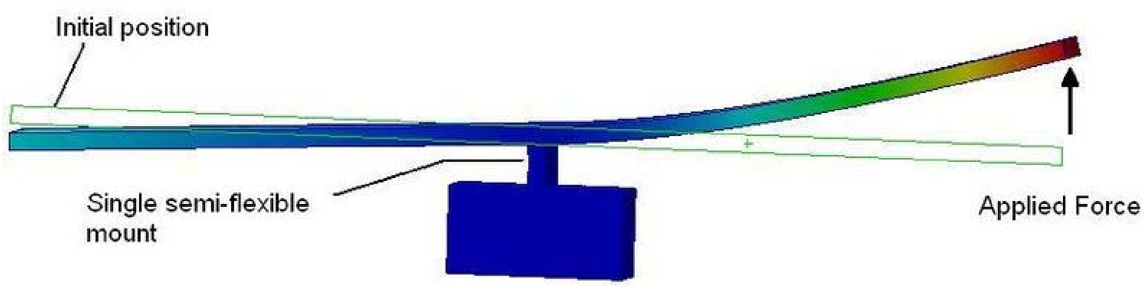
The leaf spring as an anti-roll bar

The extent to which a leaf spring acts as an anti-roll bar bar is determined by the way it is mounted. A single, loose center mount would cause the spring to pivot about the center axis, pushing one wheel down as the other was compressed upward. This is exactly opposite of an anti-roll bar and has not been used on any generation of the Corvette.

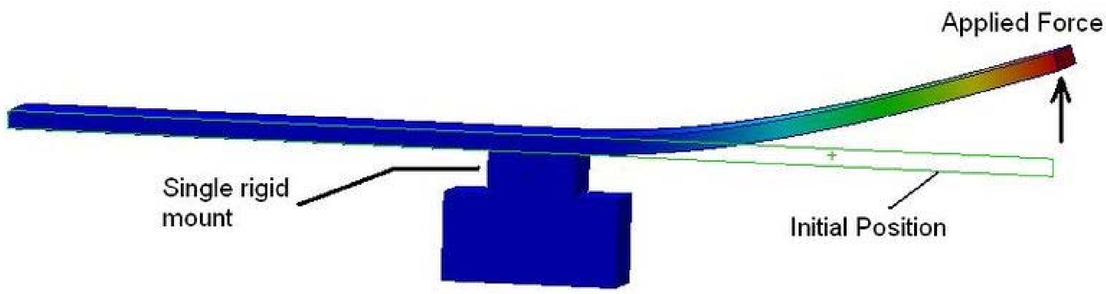
A single, perfectly rigid center mount that held a small center section of the spring flat against the frame would isolate one side of the spring from the other. No roll or anti-roll effect would appear. The rear spring of the C2, C3, and C4 has this type of mount, which effectively divides the spring in two. It becomes a quarter-elliptic spring.



A single transverse spring with a flexible center mount. When one side is pushed up the other side moves down.

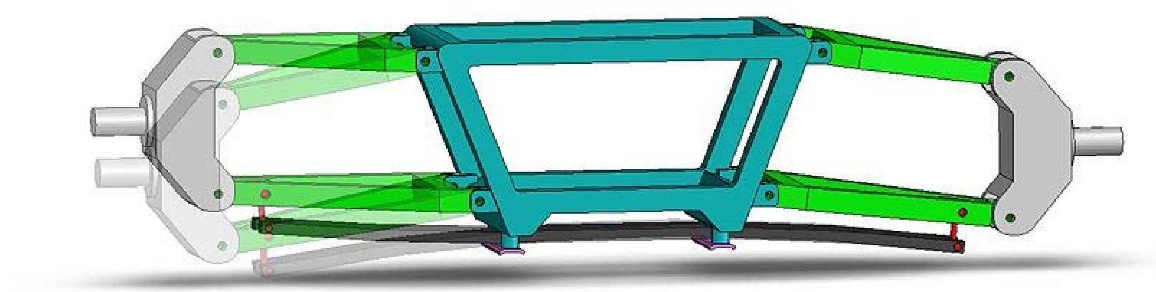


A transverse leaf spring with a semi-rigid mount. When one side is pushed up the other side moves down significantly less than in the flexible mount case.

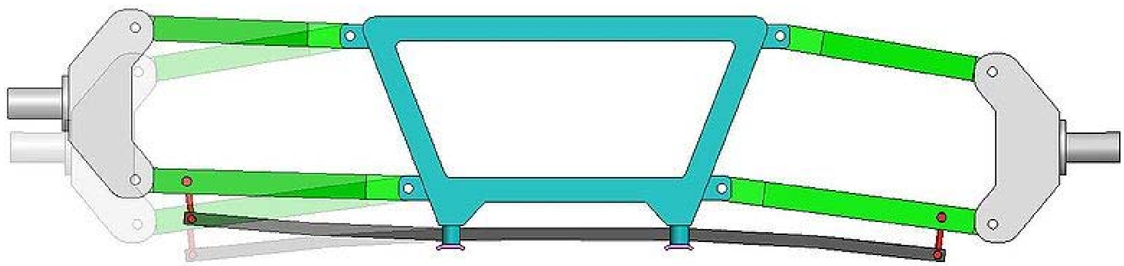


A transverse leaf spring with a central rigid mount. The two spring halves are effectively isolated. Movements of one half of the spring do not affect the other half.

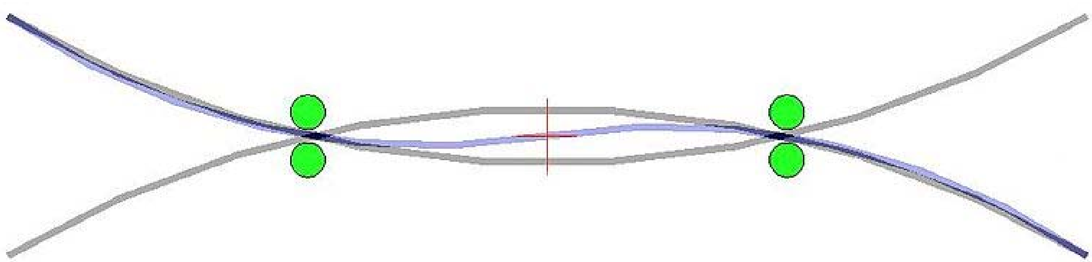
Beginning with the C4 model, the Corvette has had widely-spaced double mounts on the front. The rear spring has had double mounts since the C5. The spring is allowed to pivot about these two points. When only one wheel is compressed as in illustration #5, the portion of the spring between the mounts assumes a horizontal "S" shape. An impact that compresses the left wheel will tighten the bend radius of the right half of the spring, thereby lowering the spring rate for the right wheel like an anti-roll bar. The caster, camber, toe-in, and general orientation of the left wheel remain unchanged.



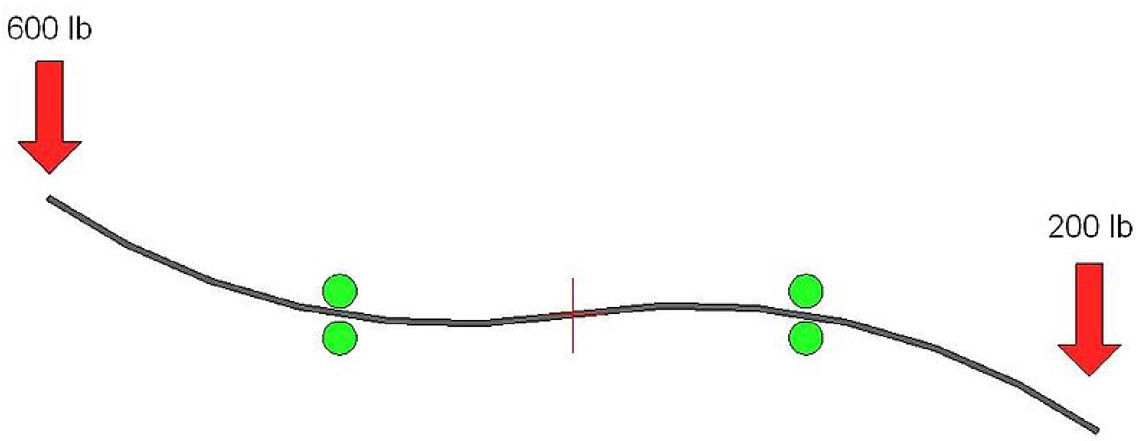
5 - The single-leaf suspension with the left side in compression.



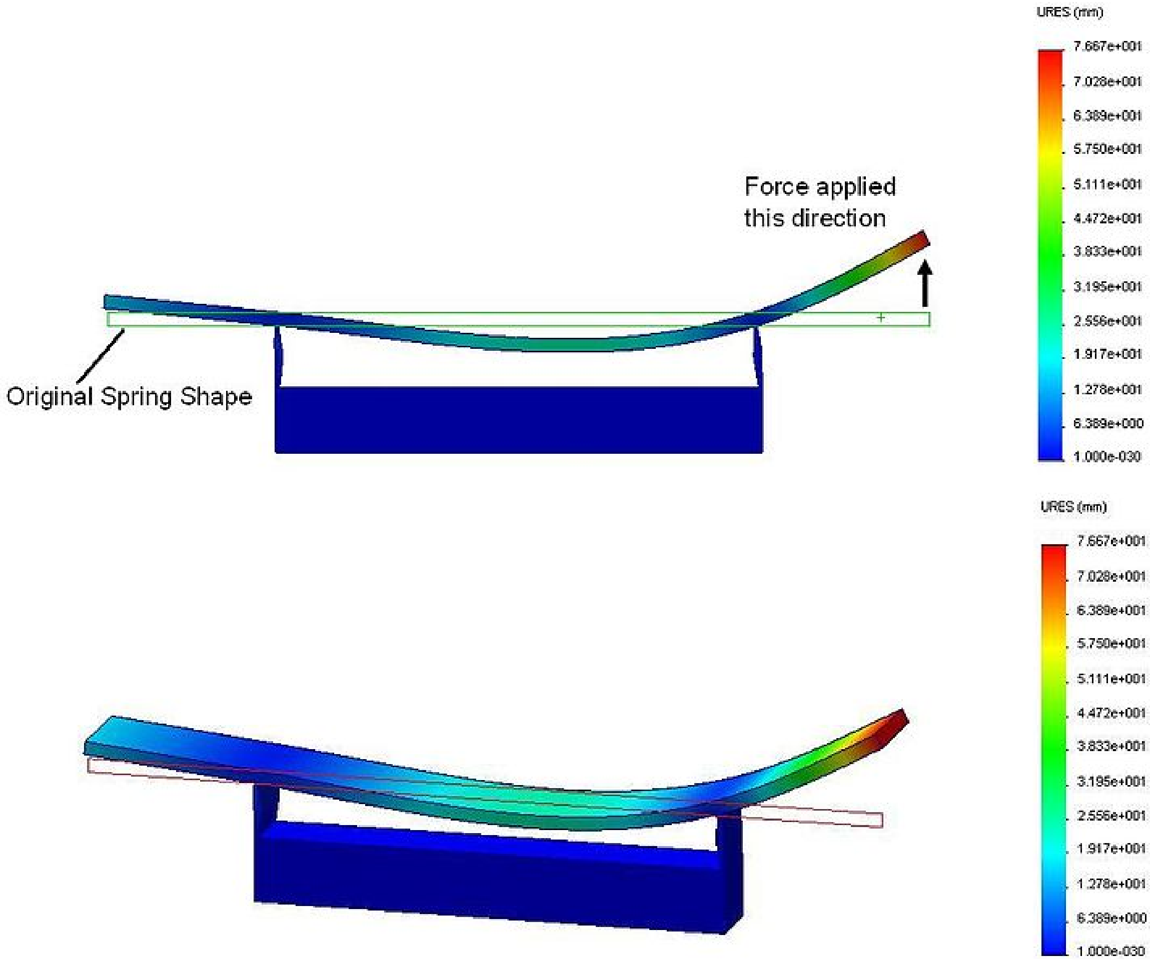
5a - The same suspension in rear profile.



This multi exposure image shows an exaggerated view of the leaf spring flex when the wheels are compressed, in droop and in roll. The S-bent spring is shown in blue.



Left side shown in compression, right side shown at static height. The left side spring force has increased from 500lbs to 600lbs while the right side has decreased from 300lb to 200lb.



Approximate FEA model of a leaf spring under load. The initial, unbent shape of the spring is shown as a silhouette box. An upward deflection on the right side of the spring results in a smaller upward movement on the left side.

With the Corvette's suspension configuration, the effects of the anti-roll bar and leaf spring add together at the wheels. This additive property allows Corvette engineers to use a smaller, lighter anti-roll bar than the car would otherwise require if it used conventional coil springs. From Dave McLellan, chief engineer on the C4 Corvette program:

We planned to use a massive front *[roll]* bar to achieve the roll stiffness we were after. We found, however, that by spreading the body attachment of the front suspension fiberglass spring into two separate attachments 18 inches apart, we could achieve a major portion of the roll stiffness contribution of the front roll bar for free. We still used a massive front bar, but it would have been even bigger and heavier if it had not been supplemented by the leaf spring.

Transverse leaf springs within independent suspensions

Advantages

- Less unsprung weight. Coil springs contribute to unsprung weight; the less there is, the more quickly the wheel can respond at a given spring rate.
- Less weight. The C4 Corvette's composite front leaf weighed 1/3 as much as the pair of conventional coil springs it would replace. Volvo reported that the single composite leaf spring used in the rear suspension of the 960 Wagon had the same mass as just one of the two springs it replaced.
- Weight is positioned lower. Coil springs and the associated chassis hard mounts raise the center of mass of the car.
- Superior wear characteristics. The Corvette's composite leaf springs last longer than coils, though in a car as light as the Corvette, the difference is not especially significant. No composite Corvette leaf has ever been replaced due to fatigue failure, though steel leafs from 1963 to 1983 have been. As of 1980, the composite spring was an option on the C3.
- As used on the Corvette, ride height can be adjusted by changing the length of the end links connecting the leaf to the suspension arms. This allows small changes in ride height with minimal effects on the spring rate.
- Also as used on the C4 front suspension, C5, and C6 Corvettes, the leaf spring acts as an anti-roll bar, allowing for smaller and lighter bars than if the car were equipped with coil springs. As implemented on the C3 and C4 rear suspensions with a rigid central mount, the anti-roll effect does not occur.
- Packaging. As used on the C5 and later Corvettes the use of OEM coil over damper springs would have forced the chassis engineers to either vertically raise the shock towers or move them inward. In the rear this would have reduced trunk space. In the front this would have interfered with engine packaging. The use of the leaf spring allowed the spring to be placed out of the way under the chassis and while keeping the diameter of the shock absorber assembly to that of just the damper rather than damper and spring.

Disadvantages

- Packaging can be problematic; the leaf must span from one side of the car to the other. This can limit applications where the drivetrain, or another part, is in the way.
- Materials expense. Steel coils are commodity items; a single composite leaf spring costs more than two of them.
- Design complexity. Composite monoleafs allow for considerable variety in shape, thickness, and materials. They are inherently more expensive to design, particularly in performance applications.
- Cost of modification. As a result of specialized design and packaging, changing spring rates often requires a custom unit. Coil springs in various sizes and rates are available inexpensively.

- Susceptibility to damage. Engine fluids and exhaust modifications like cat-back removal might weaken or destroy composite springs over time. The leaf spring is more susceptible to heat related damage than conventional steel springs.
- Perception. Due to its association with spring-located solid axles, the leaf spring has a stigma unrelated to the spring itself.

Racing concerns

- Running stiffer springs left-to-right would require either asymmetrical spring mounts or an asymmetric spring. However, a few companies such as VBP offer kits that allow independent adjustment of spring rate and ride height at all four corners of the car.
- Regulations often prohibit the use of leaf springs; NASCAR does not allow them.
- The more compact shape of a coil spring can allow for variation in more suspension design and spring placement. Because a transverse leaf spring must span the width of the car, open-wheel cars are too low to use them. The leaf spring would have to pass through the gearbox or the driver's legs.
- Coil springs are not car-specific. A Porsche, an LMP, and a Ferrari can all use a spring custom wound on the same generic equipment. Custom composite leaf springs require expensive retooling and cannot be used across car models.
- The characteristics of coil springs in a performance environment are known, and racers will use what they know. Most race teams do not have adequate experience with leaf springs to use them in this capacity.

Carroll Smith is quoted in his book, *Engineer to Win*

If I were involved in the design of a new passenger vehicle, however, I would give serious consideration to the use of a transverse composite single leaf spring of unidirectional glass or carbon filament in an epoxy matrix. This would be the lightest practical spring configuration and, although space constraints would seem to limit its use in racing, it should be perfectly feasible on road-going vehicles, from large trucks to small commuter cars. (Since I wrote this paragraph the new-generation Corvette has come out with just such a spring to control its independent suspension systems-at both end of the car.)

Transverse leaf springs in other vehicles

In addition to the Corvette, a composite transverse leaf spring has been used on other GM and non-GM vehicles.

- Volvo 960 (Wagon only)
- Volvo S90
- Mercedes Sprinter vans (transverse in front only)
- VW 1-Litre-Car prototype car
- GM W platform cars- (Lumina, Grand Prix, Regal, Cutlass Supreme).
- GM E platform cars- (Eldorado, Toronado, Riviera, Reatta, Allante).
- Mercedes Smart ForTwo (used with MacPherson Struts)

- Indigo, a Swedish made, low volume roadster. Due to the anti-roll properties of the transverse leaf spring setup the car does not use a separate front anti-roll bar.

Many small European cars such as the Fiat 128, the Yugo, and the Triumph Motor Company small chassis cars (Herald, Vitesse, Spitfire, GT6) used transverse steel springs in similar fashion. The Yugo's steel spring used twin attachment points and did provide anti-roll capability.

Recent patents and research utilizing dual pivotally supported composite leaf springs

In addition to the vehicles mentioned above, several automotive companies have researched suspension designs using a transverse composite leaf spring supported in a fashion similar to that of the Corvette.

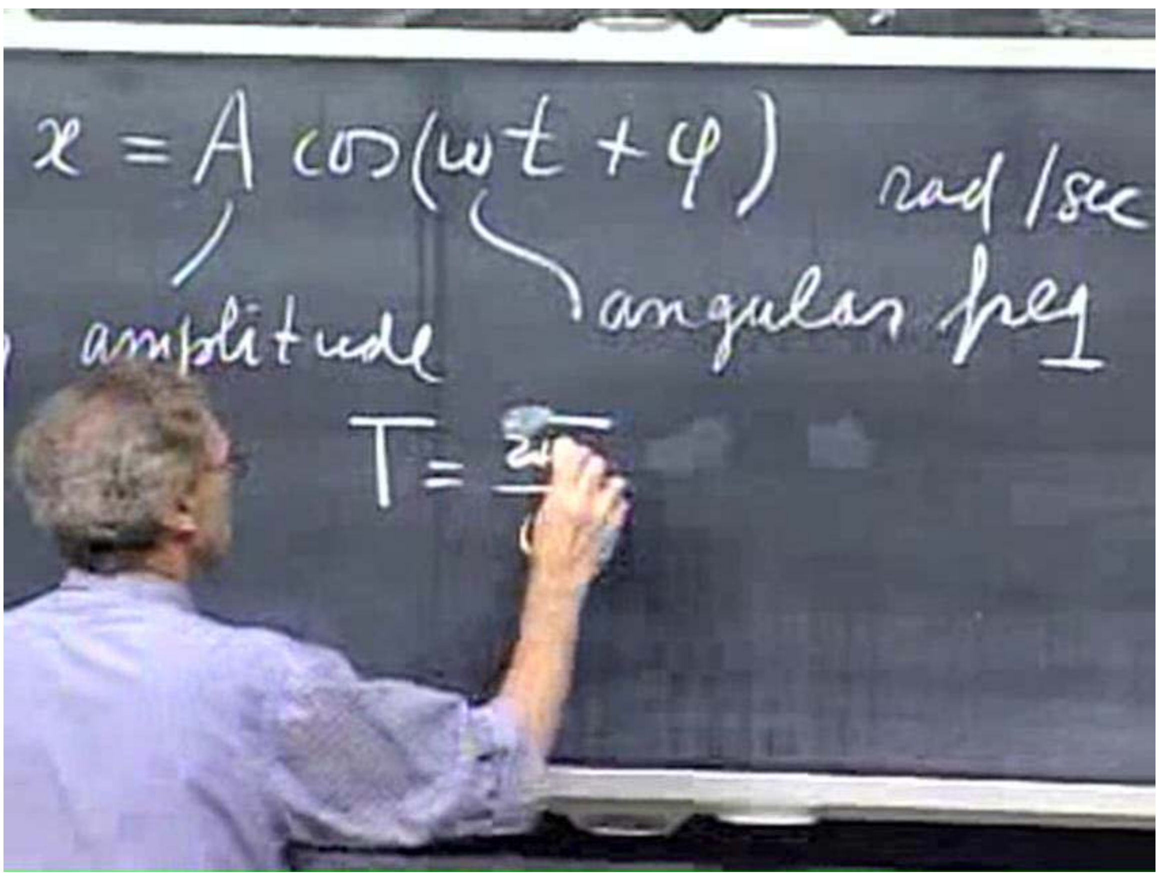
- Ford Global Technologies, 2006 patent #7029017, *Wheel suspension for a motor vehicle with a transverse leaf spring*.
- Porsche AG, 2000 patent # 6029987, *Front Axle for a Motor Vehicle*. Describes a strut suspension system supported by a transverse leaf spring system largely the same as that used by the Corvette. The Porsche patent mentions the beneficial stability effects of this arrangement
- Honda, 1992 *Transverse leaf spring type suspension* patent #5141209
- DaimlerChrysler, 2004, patent #6811169, *Composite Spring Design that also Performs the Lower Control Arm Function for a Conventional or Active Suspension System*
- ZF released a concept rear suspension design in October 2009 using a composite spring based rear suspension. The strut based suspension uses a transverse leaf spring to function as both ride and anti-roll spring. The ZF concept differs from the system used on the Corvette by using the leaf spring as one of the suspension links.

Chapter 9

Hooke's Law

Hooke's law models the properties of springs for small changes in length

Prof. Walter Lewin explains Hooke's law, "arguably the most important equation in physics", in MIT's Course 8.01



A test of Hooke's law

In mechanics, and physics, **Hooke's law** of elasticity is an approximation that states that the extension of a spring is in direct proportion with the load applied to it. Many materials obey this law as long as the load does not exceed the material's elastic limit. Materials for which Hooke's law is a useful approximation are known as linear-elastic or "Hookean" materials. Hooke's law in simple terms says that strain is directly proportional to stress.

Mathematically, Hooke's law states that

$$\mathbf{F} = -k\mathbf{x},$$

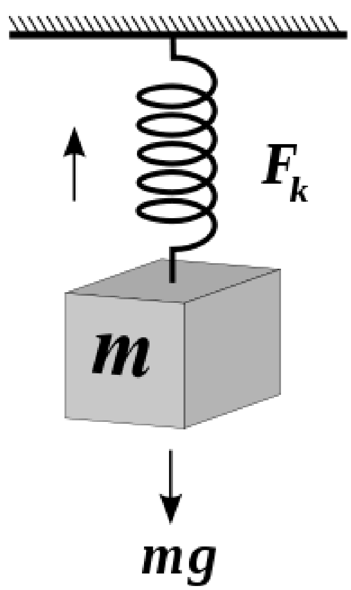
where

> ***x*** is the displacement of the spring's end from its equilibrium position (a distance, in SI units: m);
> ***F*** is the restoring force exerted by the material (in SI units: N or $kgms^{-2}$); and
> *k* is a constant called the *rate* or *spring constant* (in SI units: $N{\cdot}m^{-1}$ or kgs^{-2}).

When this holds, the behavior is said to be *linear*. If shown on a graph, the line should show a direct variation. There is a negative sign on the right hand side of the equation because the restoring force always acts in the opposite direction of the displacement (for example, when a spring is stretched to the left, it pulls back to the right).

Hooke's law is named after the 17th century British physicist Robert Hooke. He first stated this law in 1660 as a Latin anagram, whose solution he published in 1678 as *Ut tensio, sic vis*, meaning, "As the extension, so the force".

General application to elastic materials



Hooke's law describes how far the spring will stretch under a specific force

Objects that quickly regain their original shape after being deformed by a force, with the molecules or atoms of their material returning to the initial state of stable equilibrium, often obey Hooke's law.

We may view a rod of any elastic material as a linear spring. The rod has length *L* and cross-sectional area *A*. Its extension (strain) is linearly proportional to its tensile stress *σ*, by a constant factor, the inverse of its modulus of elasticity, *E*, hence,

$$\sigma = E\varepsilon$$

or

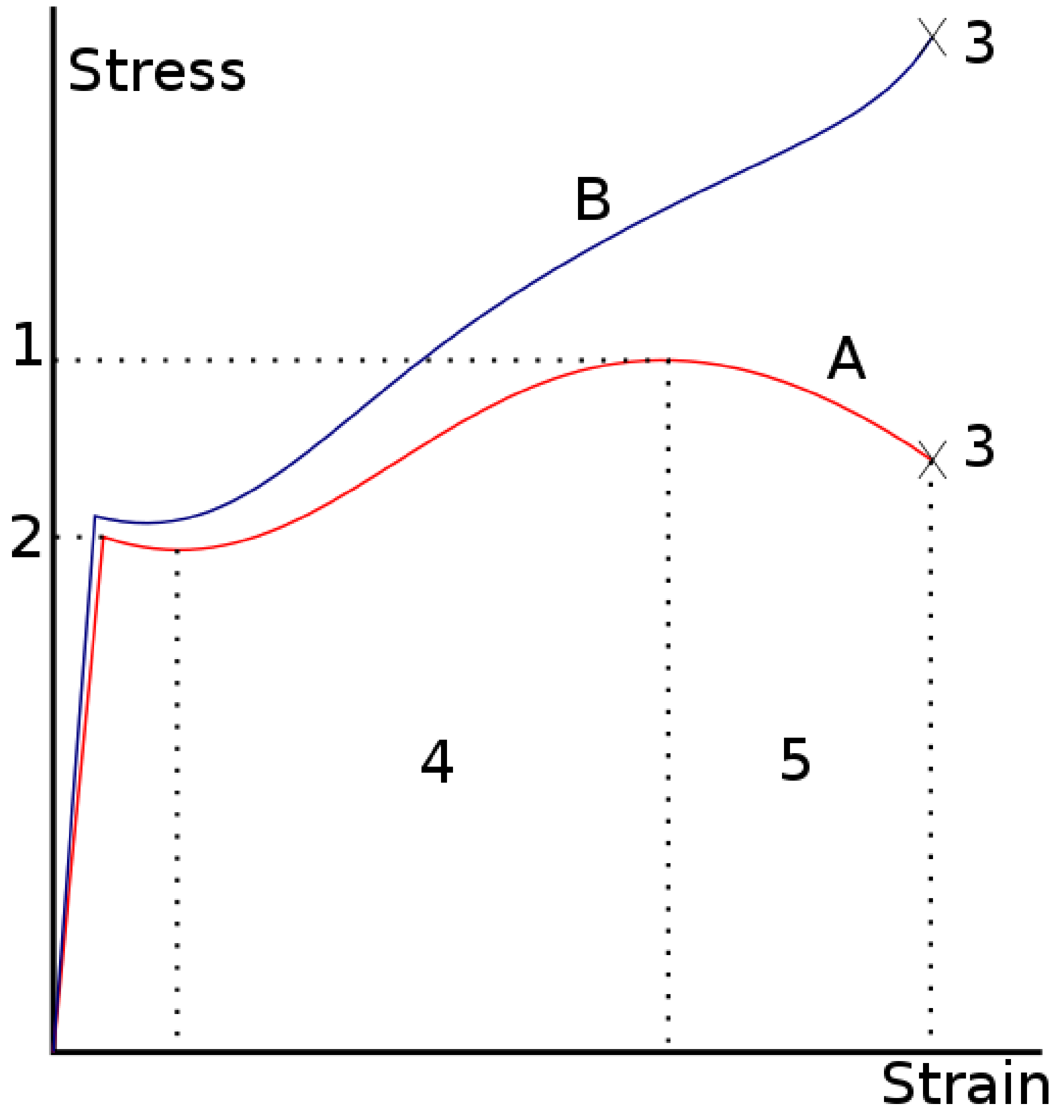
$$\Delta L = \frac{F}{EA}L = \frac{\sigma}{E}L.$$

Hooke's law only holds for some materials under certain loading conditions. Steel exhibits linear-elastic behavior in most engineering applications; Hooke's law is valid for it throughout its **elastic range** (i.e., for stresses below the yield strength). For some other materials, such as aluminium, Hooke's law is only valid for a portion of the elastic range. For these materials a proportional limit stress is defined, below which the errors associated with the linear approximation are negligible.

Rubber is generally regarded as a "non-hookean" material because its elasticity is stress dependent and sensitive to temperature and loading rate.

Applications of the law include spring operated weighing machines, stress analysis and modelling of materials.

The spring equation



Stress–strain curve for low-carbon steel. Hooke's law is only valid for the portion of the curve between the origin and the yield point(2).

1. Ultimate strength
2. Yield strength - corresponds to yield point
3. Rupture
4. Strain hardening region
5. Necking region

A: Apparent stress (F/A_0)
B: True stress (F/A)

The most commonly encountered form of Hooke's law is probably the *spring equation*, which relates the force exerted by a spring to the distance it is stretched by a *spring constant*, *k*, measured in force per length.

$$F = -kx$$

The negative sign indicates that the force exerted by the spring is in direct opposition to the direction of displacement. It is called a "restoring force", as it tends to restore the system to equilibrium. The potential energy stored in a spring is given by

$$PE = \frac{1}{2}kx^2$$

which comes from adding up the energy it takes to incrementally compress the spring. That is, the integral of force over distance. (Note that potential energy of a spring is always non-negative.)

This potential can be visualized as a parabola on the *U*-*x* plane. As the spring is stretched in the positive x-direction, the potential energy increases (the same thing happens as the spring is compressed). The corresponding point on the potential energy curve is higher than that corresponding to the equilibrium position ($x = 0$). The tendency for the spring is to therefore decrease its potential energy by returning to its equilibrium (unstretched) position, just as a ball rolls downhill to decrease its gravitational potential energy.

If a mass *m* is attached to the end of such a spring, the system becomes a harmonic oscillator. It will oscillate with a natural frequency given either as an angular frequency

$$\omega = \sqrt{\frac{k}{m}}$$

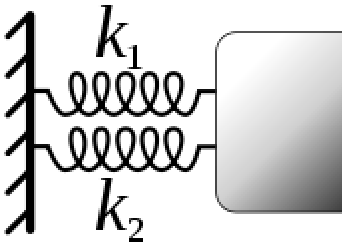
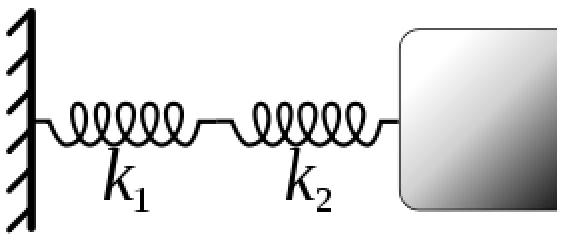
or as a natural frequency

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}}.$$

This idealized description of spring mechanics works as long as the mass of the spring is very small compared to the mass *m*, there is no significant friction on the system, and the spring is not overextended beyond its natural range (which can deform it permanently).

Multiple springs

When two springs are attached to a mass and compressed, the following table compares values of the springs.

Comparison	In Parallel	In Series
		
Equivalent spring constant	$k_{eq} = k_1 + k_2$	$\frac{1}{k_{eq}} = \frac{1}{k_1} + \frac{1}{k_2}$
Compressed distance	$x_1 = x_2$	$\frac{x_1}{x_2 - x_1} = \frac{k_2}{k_1}$
Energy stored	$\frac{E_1}{E_2} = \frac{k_1}{k_2}$	$\frac{E_1}{E_2} = \frac{k_2}{k_1}$

Derivation

Equivalent Spring Constant (Series)

$$\begin{aligned} F_b &= -k_2 x_2 + k_2 x_1 \\ &= -k_2 x_2 + k_2 \left(\frac{k_2}{k_1 + k_2} x_2 \right) \\ &= -k_2 x_2 \left(\frac{k_1 + k_2}{k_1 + k_2} \right) + \frac{k_2^2}{k_1 + k_2} x_2 \\ &= x_2 \frac{-k_1 k_2 - k_2^2 + k_2^2}{k_1 + k_2} \end{aligned}$$

Equivalent Spring Constant (Parallel)

$$\begin{aligned} F_b &= F_1 + F_2 \\ &= -k_1 x - k_2 x \end{aligned}$$

Compressed Distance

$$|F_1| = |F_2|$$

$$k_1 x_1 = k_2 (x_2 - x_1) \, .$$

Tensor expression of Hooke's Law

Note: the Einstein summation convention of summing on repeated indices is used below.

When working with a three-dimensional stress state, a 4th order tensor c($c_{ijk\ell}$) containing 81 elastic coefficients must be defined to link the stress tensor $\boldsymbol{\sigma}$(σ_{ij}) and the strain tensor $\boldsymbol{\epsilon}$($\epsilon_{k\ell}$).

$$\boldsymbol{\sigma} = \mathsf{c} : \boldsymbol{\epsilon} \, .$$

Expressed in terms of components with respect to an orthonormal basis, the generalized form of Hooke's law is written as (using the summation convention)

$$\sigma_{ij} = c_{ijk\ell} \, \epsilon_{k\ell}$$

The tensor c is called the **stiffness tensor** or the **elasticity tensor**. Due to the symmetry of the stress tensor, strain tensor, and stiffness tensor, only 21 elastic coefficients are independent. As stress is measured in units of pressure and strain is dimensionless, the entries of $c_{ijk\ell}$ are also in units of pressure.

The expression for generalized Hooke's law can be inverted to get a relation for the strain in terms of stress:

$$\boldsymbol{\epsilon} = \mathsf{s} : \boldsymbol{\sigma} \qquad \text{or} \qquad \epsilon_{ij} = s_{ijk\ell} \, \sigma_{k\ell} \, .$$

The tensor s is called the **compliance tensor**.

Generalization for the case of large deformations is provided by models of neo-Hookean solids and Mooney-Rivlin solids.

Isotropic materials

Isotropic materials are characterized by properties which are independent of direction in space. Physical equations involving isotropic materials must therefore be independent of the coordinate system chosen to represent them. The strain tensor is a symmetric tensor. Since the trace of any tensor is independent of any coordinate system, the most complete coordinate-free decomposition of a symmetric tensor is to represent it as the sum of a constant tensor and a traceless symmetric tensor.[:Ch. 10] Thus:

$$\varepsilon_{ij} = \left(\tfrac{1}{3}\varepsilon_{kk}\delta_{ij}\right) + \left(\varepsilon_{ij} - \tfrac{1}{3}\varepsilon_{kk}\delta_{ij}\right)$$

where δ_{ij} is the Kronecker delta. In direct tensor notation

$$\boldsymbol{\varepsilon} = \text{vol}(\boldsymbol{\varepsilon}) + \text{dev}(\boldsymbol{\varepsilon}) \ ; \quad \text{vol}(\boldsymbol{\varepsilon}) := \tfrac{1}{3}\,\text{tr}(\boldsymbol{\varepsilon})\,\mathbf{I} \ ; \quad \text{dev}(\boldsymbol{\varepsilon}) := \boldsymbol{\varepsilon} - \text{vol}(\boldsymbol{\varepsilon})$$

where $\mathbf{I}$ is the second-order identity tensor. The first term on the right is the constant tensor, also known as the **volumetric strain tensor**, and the second term is the traceless symmetric tensor, also known as the **deviatoric strain tensor** or shear tensor.

The most general form of Hooke's law for isotropic materials may now be written as a linear combination of these two tensors:

$$\sigma_{ij} = 3K\left(\tfrac{1}{3}\varepsilon_{kk}\delta_{ij}\right) + 2G\left(\varepsilon_{ij} - \tfrac{1}{3}\varepsilon_{kk}\delta_{ij}\right) \ ; \quad \boldsymbol{\sigma} = 3K\ \text{vol}(\boldsymbol{\varepsilon}) + 2G\ \text{dev}(\boldsymbol{\varepsilon})$$

where K is the bulk modulus and G is the shear modulus.

Using the relationships between the elastic moduli, these equations may also be expressed in various other ways. A common form of Hooke's law for isotropic materials, expressed in direct tensor notation, is

$$\boldsymbol{\sigma} = \lambda\ \text{tr}(\boldsymbol{\varepsilon})\ \mathbf{I} + 2\mu\ \boldsymbol{\varepsilon} = \mathsf{c} : \boldsymbol{\varepsilon} \ ; \quad \mathsf{c} = \lambda\ \mathbf{I}\otimes\mathbf{I} + 2\mu\ \mathsf{I}$$

where $\lambda := K - 2/3G$ and $\mu := G$ are the Lamé constants, $\mathbf{I}$ is the second-order identity tensor, and I is the symmetric part of the fourth-order identity tensor. In terms of components with respect to a Cartesian basis,

$$\sigma_{ij} = \lambda\ \varepsilon_{kk}\ \delta_{ij} + 2\mu\ \varepsilon_{ij} = c_{ijk\ell}\ \varepsilon_{k\ell} \ ; \quad c_{ijk\ell} = \lambda\ \delta_{ij}\ \delta_{k\ell} + \mu\ (\delta_{ik}\ \delta_{j\ell} + \delta_{i\ell}\ \delta_{jk})$$

The inverse relationship is

$$\boldsymbol{\varepsilon} = \frac{1}{2\mu}\ \boldsymbol{\sigma} - \frac{\lambda}{2\mu(3\lambda+2\mu)}\ \text{tr}(\boldsymbol{\sigma})\ \mathbf{I} = \frac{1}{2G}\ \boldsymbol{\sigma} + \left(\frac{1}{9K} - \frac{1}{6G}\right)\ \text{tr}(\boldsymbol{\sigma})\ \mathbf{I}$$

Therefore the compliance tensor in the relation $\boldsymbol{\varepsilon} = \mathsf{s} : \boldsymbol{\sigma}$ is

$$\mathsf{s} = -\frac{\lambda}{2\mu(3\lambda+2\mu)}\ \mathbf{I}\otimes\mathbf{I} + \frac{1}{2\mu}\ \mathsf{I} = \left(\frac{1}{9K} - \frac{1}{6G}\right)\ \mathbf{I}\otimes\mathbf{I} + \frac{1}{2G}\ \mathsf{I}$$

In terms of Young's modulus and Poisson's ratio, Hooke's law for isotropic materials can then be expressed as

$$\boldsymbol{\varepsilon} = \frac{1}{E}\ \boldsymbol{\sigma} - \frac{\nu}{E}\left[\text{tr}(\boldsymbol{\sigma})\ \mathbf{I} - \boldsymbol{\sigma}\right]$$

This is the form in which the strain is expressed in terms of the stress tensor in engineering. The expression in expanded form is

$$\varepsilon_{11} = \tfrac{1}{E}\left[\sigma_{11} - \nu(\sigma_{22} + \sigma_{33})\right]$$
$$\varepsilon_{22} = \tfrac{1}{E}\left[\sigma_{22} - \nu(\sigma_{11} + \sigma_{33})\right]$$
$$\varepsilon_{33} = \tfrac{1}{E}\left[\sigma_{33} - \nu(\sigma_{11} + \sigma_{22})\right]$$
$$\varepsilon_{12} = \tfrac{1}{2G}\,\sigma_{12}\ ;\quad \varepsilon_{13} = \tfrac{1}{2G}\,\sigma_{13}\ ;\quad \varepsilon_{23} = \tfrac{1}{2G}\,\sigma_{23}$$

where *E* is the modulus of elasticity and ν is Poisson's ratio.

In matrix form, Hooke's law for isotropic materials can be written as

$$\begin{bmatrix}\varepsilon_{11}\\ \varepsilon_{22}\\ \varepsilon_{33}\\ 2\varepsilon_{23}\\ 2\varepsilon_{31}\\ 2\varepsilon_{12}\end{bmatrix} = \begin{bmatrix}\varepsilon_{11}\\ \varepsilon_{22}\\ \varepsilon_{33}\\ \gamma_{23}\\ \gamma_{31}\\ \gamma_{12}\end{bmatrix} = \frac{1}{E}\begin{bmatrix}1 & -\nu & -\nu & 0 & 0 & 0\\ -\nu & 1 & -\nu & 0 & 0 & 0\\ -\nu & -\nu & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 2(1+\nu) & 0 & 0\\ 0 & 0 & 0 & 0 & 2(1+\nu) & 0\\ 0 & 0 & 0 & 0 & 0 & 2(1+\nu)\end{bmatrix}\begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{31}\\ \sigma_{12}\end{bmatrix}$$

where $\gamma_{ij} := 2\varepsilon_{ij}$ is the **engineering shear strain**. The inverse relation may be written as

$$\begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{31}\\ \sigma_{12}\end{bmatrix} = \frac{E}{(1+\nu)(1-2\nu)}\begin{bmatrix}1-\nu & \nu & \nu & 0 & 0 & 0\\ \nu & 1-\nu & \nu & 0 & 0 & 0\\ \nu & \nu & 1-\nu & 0 & 0 & 0\\ 0 & 0 & 0 & (1-2\nu)/2 & 0 & 0\\ 0 & 0 & 0 & 0 & (1-2\nu)/2 & 0\\ 0 & 0 & 0 & 0 & 0 & (1-2\nu)/2\end{bmatrix}\begin{bmatrix}\varepsilon_{11}\\ \varepsilon_{22}\\ \varepsilon_{33}\\ 2\varepsilon_{23}\\ 2\varepsilon_{31}\\ 2\varepsilon_{12}\end{bmatrix}$$

which expression can be simplified thanks to the Lamé constants :

$$\begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{31}\\ \sigma_{12}\end{bmatrix} = \begin{bmatrix}2\mu+\lambda & \lambda & \lambda & 0 & 0 & 0\\ \lambda & 2\mu+\lambda & \lambda & 0 & 0 & 0\\ \lambda & \lambda & 2\mu+\lambda & 0 & 0 & 0\\ 0 & 0 & 0 & \mu & 0 & 0\\ 0 & 0 & 0 & 0 & \mu & 0\\ 0 & 0 & 0 & 0 & 0 & \mu\end{bmatrix}\begin{bmatrix}\varepsilon_{11}\\ \varepsilon_{22}\\ \varepsilon_{33}\\ 2\varepsilon_{23}\\ 2\varepsilon_{31}\\ 2\varepsilon_{12}\end{bmatrix}$$

Plane stress Hooke's law

Under plane stress conditions $\sigma_{33} = \sigma_{31} = \sigma_{23} = 0$. In that case Hooke's law takes the form

$$\begin{bmatrix}\varepsilon_{11}\\ \varepsilon_{22}\\ \varepsilon_{33}\\ 2\varepsilon_{12}\end{bmatrix} = \frac{1}{E}\begin{bmatrix}1 & -\nu & 0\\ -\nu & 1 & 0\\ -\nu & -\nu & 0\\ 0 & 0 & 2(1+\nu)\end{bmatrix}\begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{12}\end{bmatrix}$$

The inverse relation is usually written in the reduced form

$$\begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{12}\end{bmatrix} = \frac{E}{1-\nu^2}\begin{bmatrix}1 & \nu & 0\\ \nu & 1 & 0\\ 0 & 0 & \frac{1-\nu}{2}\end{bmatrix}\begin{bmatrix}\varepsilon_{11}\\ \varepsilon_{22}\\ 2\varepsilon_{12}\end{bmatrix}$$

Anisotropic materials

The symmetry of the Cauchy stress tensor ($\sigma_{ij} = \sigma_{ji}$) and the generalized Hooke's laws ($\sigma_{ij} = c_{ijk\ell}\ \epsilon_{k\ell}$) implies that $c_{ijk\ell} = c_{jik\ell}$. Similarly, the symmetry of the infinitesimal strain tensor implies that $c_{ijk\ell} = c_{ij\ell k}$. These symmetries are called the **minor symmetries** of the **stiffness tensor** (c).

If in addition, since the displacement gradient and the Cauchy stress are work conjugate, the stress-strain relation can be derived from a strain energy density functional (*U*), then

$$\sigma_{ij} = \frac{\partial U}{\partial \epsilon_{ij}} \quad \implies \quad c_{ijk\ell} = \frac{\partial^2 U}{\partial \epsilon_{ij}\partial \epsilon_{k\ell}}\ .$$

The arbitrariness of the order of differentiation implies that $c_{ijk\ell} = c_{k\ell ij}$. These are called the **major symmetries** of the stiffness tensor. The major and minor symmetries indicate that the stiffness tensor has only 21 independent components.

Matrix representation (stiffness tensor)

It is often useful to express the anisotropic form of Hooke's law in matrix notation, also called Voigt notation. To do this we take advantage of the symmetry of the stress and strain tensors and express them as six-dimensional vectors in an orthonormal coordinate system ($\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$) as

$$[\boldsymbol{\sigma}] = \begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{31}\\ \sigma_{12}\end{bmatrix} \equiv \begin{bmatrix}\sigma_{1}\\ \sigma_{2}\\ \sigma_{3}\\ \sigma_{4}\\ \sigma_{5}\\ \sigma_{6}\end{bmatrix} ~;~~ [\boldsymbol{\epsilon}] = \begin{bmatrix}\epsilon_{11}\\ \epsilon_{22}\\ \epsilon_{33}\\ 2\epsilon_{23}\\ 2\epsilon_{31}\\ 2\epsilon_{12}\end{bmatrix} \equiv \begin{bmatrix}\epsilon_{1}\\ \epsilon_{2}\\ \epsilon_{3}\\ \epsilon_{4}\\ \epsilon_{5}\\ \epsilon_{6}\end{bmatrix}$$

Then the stiffness tensor (C) can be expressed as

$$[\mathsf{C}] = \begin{bmatrix} c_{1111} & c_{1122} & c_{1133} & c_{1123} & c_{1131} & c_{1112} \\ c_{2211} & c_{2222} & c_{2233} & c_{2223} & c_{2231} & c_{2212} \\ c_{3311} & c_{3322} & c_{3333} & c_{3323} & c_{3331} & c_{3312} \\ c_{2311} & c_{2322} & c_{2333} & c_{2323} & c_{2331} & c_{2312} \\ c_{3111} & c_{3122} & c_{3133} & c_{3123} & c_{3131} & c_{3112} \\ c_{1211} & c_{1222} & c_{1233} & c_{1223} & c_{1231} & c_{1212} \end{bmatrix} \equiv \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix}$$

and Hooke's law is written as

$$[\boldsymbol{\sigma}] = [\mathsf{C}][\boldsymbol{\epsilon}] \qquad \text{or} \qquad \sigma_i = C_{ij}\epsilon_j ~.$$

Similarly the compliance tensor (S) can be written as

$$[\mathsf{S}] = \begin{bmatrix} s_{1111} & s_{1122} & s_{1133} & 2s_{1123} & 2s_{1131} & 2s_{1112} \\ s_{2211} & s_{2222} & s_{2233} & 2s_{2223} & 2s_{2231} & 2s_{2212} \\ s_{3311} & s_{3322} & s_{3333} & 2s_{3323} & 2s_{3331} & 2s_{3312} \\ 2s_{2311} & 2s_{2322} & 2s_{2333} & 4s_{2323} & 4s_{2331} & 4s_{2312} \\ 2s_{3111} & 2s_{3122} & 2s_{3133} & 4s_{3123} & 4s_{3131} & 4s_{3112} \\ 2s_{1211} & 2s_{1222} & 2s_{1233} & 4s_{1223} & 4s_{1231} & 4s_{1212} \end{bmatrix} \equiv \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16} \\ S_{12} & S_{22} & S_{23} & S_{24} & S_{25} & S_{26} \\ S_{13} & S_{23} & S_{33} & S_{34} & S_{35} & S_{36} \\ S_{14} & S_{24} & S_{34} & S_{44} & S_{45} & S_{46} \\ S_{15} & S_{25} & S_{35} & S_{45} & S_{55} & S_{56} \\ S_{16} & S_{26} & S_{36} & S_{46} & S_{56} & S_{66} \end{bmatrix}$$

Change of coordinate system

If a linear elastic material is rotated from a reference configuration to another, then the material is symmetric with respect to the rotation if the components of the stiffness tensor in the rotated configuration are related to the components in the reference configuration by the relation

$$c_{pqrs} = l_{pi}~l_{qj}~l_{rk}~l_{s\ell}~c_{ijk\ell}$$

where l_{ab} are the components of an orthogonal rotation matrix $[L]$. The same relation also holds for inversions.

In matrix notation, if the transformed basis (rotated or inverted) is related to the reference basis by

$$[\mathbf{e}_i'] = [L][\mathbf{e}_i]$$

then

$$C_{ij}\ \epsilon_i\ \epsilon_j = C'_{ij}\ \epsilon'_i\ \epsilon'_j\ .$$

In addition, if the material is symmetric with respect to the transformation [L] then

$$C_{ij} = C'_{ij} \quad \implies \quad C_{ij}\ (\epsilon_i\ \epsilon_j - \epsilon'_i\ \epsilon'_j) = 0\ .$$

Orthotropic materials

Orthotropic materials have three orthogonal planes of symmetry. If the basis vectors ($\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$) are normals to the planes of symmetry then the coordinate transformation relations imply that

$$\begin{bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6\end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0\\ C_{12} & C_{22} & C_{23} & 0 & 0 & 0\\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0\\ 0 & 0 & 0 & C_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & C_{55} & 0\\ 0 & 0 & 0 & 0 & 0 & C_{66}\end{bmatrix}\begin{bmatrix}\epsilon_1\\ \epsilon_2\\ \epsilon_3\\ \epsilon_4\\ \epsilon_5\\ \epsilon_6\end{bmatrix}$$

The inverse of this relation is commonly written as

$$\begin{bmatrix}\epsilon_{\rm xx}\\ \epsilon_{\rm yy}\\ \epsilon_{\rm zz}\\ 2\epsilon_{\rm yz}\\ 2\epsilon_{\rm zx}\\ 2\epsilon_{\rm xy}\end{bmatrix} = \begin{bmatrix} \frac{1}{E_{\rm x}} & -\frac{\nu_{\rm xy}}{E_{\rm x}} & -\frac{\nu_{\rm xz}}{E_{\rm x}} & 0 & 0 & 0\\ -\frac{\nu_{\rm yx}}{E_{\rm y}} & \frac{1}{E_{\rm y}} & -\frac{\nu_{\rm yz}}{E_{\rm y}} & 0 & 0 & 0\\ -\frac{\nu_{\rm zx}}{E_{\rm z}} & -\frac{\nu_{\rm zy}}{E_{\rm z}} & \frac{1}{E_{\rm z}} & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{1}{G_{\rm yz}} & 0 & 0\\ 0 & 0 & 0 & 0 & \frac{1}{G_{\rm zx}} & 0\\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G_{\rm xy}}\end{bmatrix}\begin{bmatrix}\sigma_{\rm xx}\\ \sigma_{\rm yy}\\ \sigma_{\rm zz}\\ \sigma_{\rm yz}\\ \sigma_{\rm zx}\\ \sigma_{\rm xy}\end{bmatrix}$$

where

$E_{\rm i}$ is the Young's modulus along axis *i*
$G_{\rm ij}$ is the shear modulus in direction *j* on the plane whose normal is in direction *i*
$\nu_{\rm ij}$ is the Poisson's ratio that corresponds to a contraction in direction *j* when an extension is applied in direction *i*.

Transversely isotropic materials

A transversely isotropic material is symmetric with respect to a rotation about an axis of symmetry. For such a material, if $\mathbf{e}_3$ is the axis of symmetry, Hooke's law can be expressed as

$$\begin{bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6\end{bmatrix} = \begin{bmatrix}C_{11} & C_{12} & C_{13} & 0 & 0 & 0\\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0\\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0\\ 0 & 0 & 0 & C_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & C_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(C_{11}-C_{12})\end{bmatrix}\begin{bmatrix}\epsilon_1\\ \epsilon_2\\ \epsilon_3\\ \epsilon_4\\ \epsilon_5\\ \epsilon_6\end{bmatrix}$$

More frequently, the $x \equiv \mathbf{e}_1$ axis is taken to be the axis of symmetry and the inverse Hooke's law is written as

$$\begin{bmatrix}\epsilon_{\rm xx}\\ \epsilon_{\rm yy}\\ \epsilon_{\rm zz}\\ 2\epsilon_{\rm yz}\\ 2\epsilon_{\rm zx}\\ 2\epsilon_{\rm xy}\end{bmatrix} = \begin{bmatrix}\frac{1}{E_{\rm x}} & -\frac{\nu_{\rm xy}}{E_{\rm x}} & -\frac{\nu_{\rm xy}}{E_{\rm x}} & 0 & 0 & 0\\ -\frac{\nu_{\rm yx}}{E_{\rm y}} & \frac{1}{E_{\rm y}} & -\frac{\nu_{\rm yz}}{E_{\rm y}} & 0 & 0 & 0\\ -\frac{\nu_{\rm yx}}{E_{\rm y}} & -\frac{\nu_{\rm zy}}{E_{\rm y}} & \frac{1}{E_{\rm y}} & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{2(1+\nu_{\rm yz})}{E_{\rm y}} & 0 & 0\\ 0 & 0 & 0 & 0 & \frac{1}{G_{\rm xy}} & 0\\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G_{\rm xy}}\end{bmatrix}\begin{bmatrix}\sigma_{\rm xx}\\ \sigma_{\rm yy}\\ \sigma_{\rm zz}\\ \sigma_{\rm yz}\\ \sigma_{\rm zx}\\ \sigma_{\rm xy}\end{bmatrix}$$

Thermodynamic basis of Hooke's law

Linear deformations of elastic materials can be approximated as adiabatic. Under these conditions and for quasistatic processes the first law of thermodynamics for a deformed body can be expressed as

$$\delta W = \delta U$$

where δU is the increase in internal energy and δW is the work done by external forces. The work can be split into two terms

$$\delta W = \delta W_s + \delta W_b$$

where δW_s is the work done by surface forces while δW_b is the work done by body forces. If $\delta\mathbf{u}$ is a variation of the displacement field $\mathbf{u}$ in the body, then the two external work terms can be expressed as

$$\delta W_s = \int_{\partial\Omega} \mathbf{t} \cdot \delta\mathbf{u}\ \mathrm{dS}\ ; \quad \delta W_b = \int_{\Omega} \mathbf{b} \cdot \delta\mathbf{u}\ \mathrm{dV}$$

where $\mathbf{t}$ is the surface traction vector, $\mathbf{b}$ is the body force vector, Ω represents the body and $\partial\Omega$ represents its surface. Using the relation between the Cauchy stress and the surface traction, $\mathbf{t} = \mathbf{n} \cdot \boldsymbol{\sigma}$ (where $\mathbf{n}$ is the unit outward normal to $\partial\Omega$), we have

$$\delta W = \delta U = \int_{\partial\Omega} (\mathbf{n} \cdot \boldsymbol{\sigma}) \cdot \delta\mathbf{u}\ \mathrm{dS} + \int_{\Omega} \mathbf{b} \cdot \delta\mathbf{u}\ \mathrm{dV}$$

Converting the surface integral into a volume integral via the divergence theorem gives

$$\delta U = \int_{\Omega} [\boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \cdot \delta\mathbf{u}) + \mathbf{b} \cdot \delta\mathbf{u}]\ \mathrm{dV}\ .$$

Using the symmetry of the Cauchy stress and the identity

$$\boldsymbol{\nabla} \cdot (\boldsymbol{A} \cdot \mathbf{b}) = (\boldsymbol{\nabla} \cdot \boldsymbol{A}) \cdot \mathbf{b} + \tfrac{1}{2}[\boldsymbol{A}^T : \boldsymbol{\nabla}\mathbf{b} + \boldsymbol{A} : (\boldsymbol{\nabla}\mathbf{b})^T]$$

we have

$$\delta U = \int_{\Omega} [\boldsymbol{\sigma} : \tfrac{1}{2}\{\boldsymbol{\nabla}\delta\mathbf{u} + (\boldsymbol{\nabla}\delta\mathbf{u})^T\} + \{\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} + \mathbf{b}\} \cdot \delta\mathbf{u}]\ \mathrm{dV}\ .$$

From the definition of strain and from the equations of equilibrium we have

$$\delta\boldsymbol{\epsilon} = \tfrac{1}{2}[\boldsymbol{\nabla}\delta\mathbf{u} + (\boldsymbol{\nabla}\delta\mathbf{u})^T]\ ; \quad \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} + \mathbf{b} = \mathbf{0}\ .$$

Hence we can write

$$\delta U = \int_{\Omega} \boldsymbol{\sigma} : \delta\boldsymbol{\epsilon}\ \mathrm{dV}$$

and therefore the variation in the internal energy density is given by

$$\delta U_0 = \boldsymbol{\sigma} : \delta\boldsymbol{\epsilon}\ .$$

An elastic material is defined as one in which the total internal energy is equal to the potential energy of the internal forces (also called the **elastic strain energy**). Therefore the internal energy density is a function of the strains, $U_0 = U_0(\boldsymbol{\epsilon})$ and the variation of the internal energy can be expressed as

$$\delta U_0 = \frac{\partial U_0}{\partial \boldsymbol{\epsilon}} : \delta \boldsymbol{\epsilon} \ .$$

Since the variation of strain is arbitrary, the stress-strain relation of an elastic material is given by

$$\boldsymbol{\sigma} = \frac{\partial U_0}{\partial \boldsymbol{\epsilon}} \ .$$

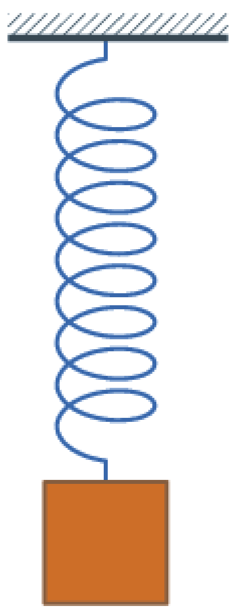
For a linear elastic material, the quantity $\partial U_0 / \partial \boldsymbol{\epsilon}$ is a linear function of $\boldsymbol{\epsilon}$, and can therefore be expressed as

$$\boldsymbol{\sigma} = \mathsf{c} : \boldsymbol{\epsilon}$$

where c is a fourth-order tensor of material constants, also called the **stiffness tensor**.

Chapter 10

Harmonic Oscillator



An undamped spring-mass system is a simple harmonic oscillator.

In classical mechanics, a **harmonic oscillator** is a system that, when displaced from its equilibrium position, experiences a restoring force, *F*, proportional to the displacement, *x*:

$$F = -kx$$

where k is a positive constant.

If F is the only force acting on the system, the system is called a **simple harmonic oscillator**, and it undergoes simple harmonic motion: sinusoidal oscillations about the equilibrium point, with a constant amplitude and a constant frequency (which does not depend on the amplitude).

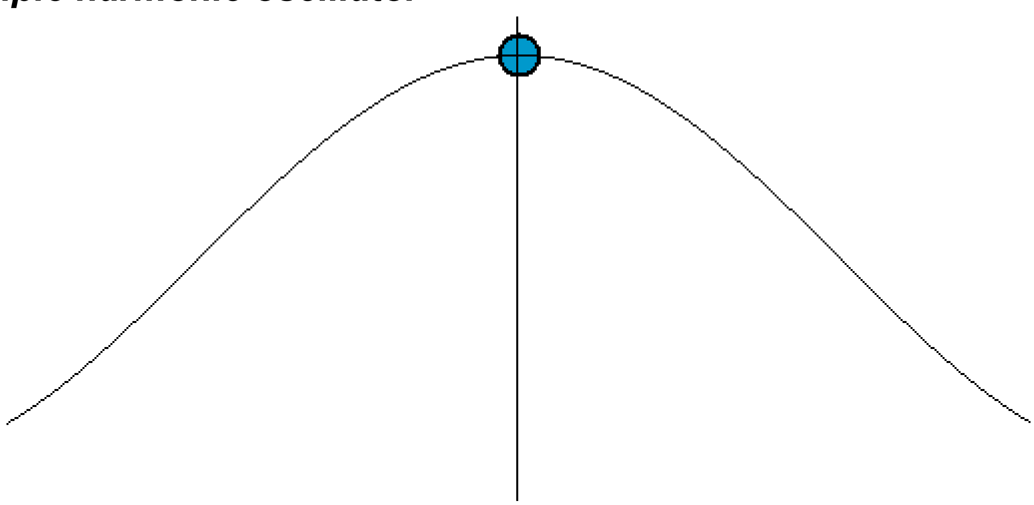
If a frictional force (damping) proportional to the velocity is also present, the harmonic oscillator is described as a **damped oscillator**. Depending on the friction coefficient, the system can:

- Oscillate with a frequency smaller than in the non-damped case, and an amplitude decreasing with time (*underdamped* oscillator).
- Decay exponentially to the equilibrium position, without oscillations (*overdamped* oscillator).

If an external time dependent force is present, the harmonic oscillator is described as a **driven oscillator**.

Mechanical examples include pendula (with small angles of displacement), masses connected to springs, and acoustical systems. Other analogous systems include electrical harmonic oscillators such as RLC circuits. The harmonic oscillator model is very important in physics, because any mass subject to a force in stable equilibrium acts as a harmonic oscillator for small vibrations. Harmonic oscillators occur widely in nature and are exploited in many manmade devices, such as clocks and radio circuits. They are the source of virtually all sinusoidal vibrations and waves.

Simple harmonic oscillator



Simple harmonic motion.

A simple harmonic oscillator is an oscillator that is neither driven nor damped. It consists of a mass m, which experiences a single force, F, which pulls the mass in the direction of the point x=0 and depends only on the mass's position x and a constant k. Newton's second law for the system is

$$F = ma = m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx.$$

Solving this differential equation, we find that the motion is described by the function

$$x(t) = A\cos\left(2\pi ft + \phi\right),$$

where

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}} = \frac{1}{T}.$$

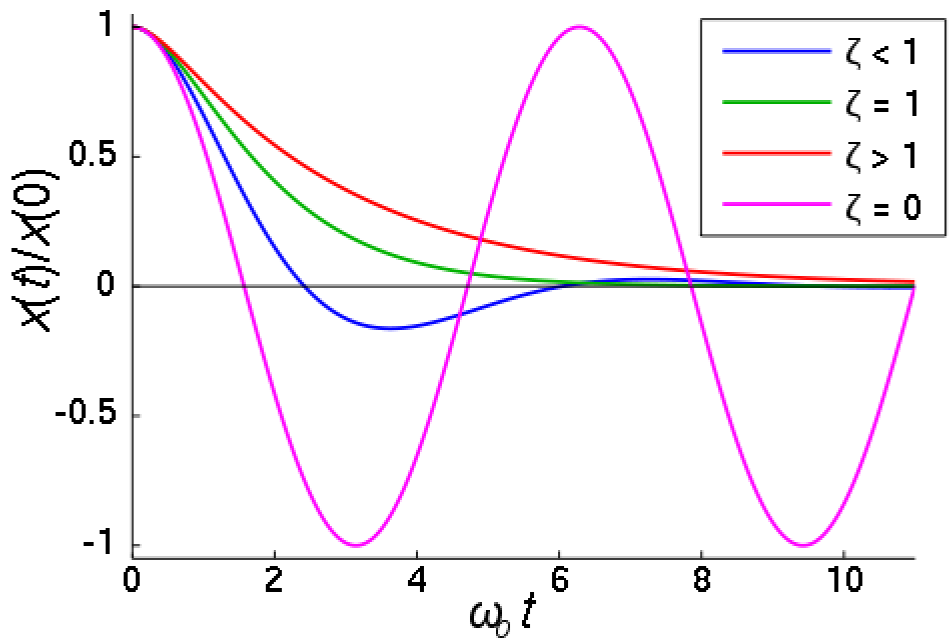
The motion is periodic— repeating itself in a sinusoidal fashion with constant amplitude, A. In addition to its amplitude, the motion of a simple harmonic oscillator is characterized by its period T, the time for a single oscillation or its frequency $f = {}^1\!/_T$, the number of cycles per unit time. The position at a given time t also depends on the phase, φ, which determines the starting point on the sine wave. The period and frequency are determined by the size of the mass m and the force constant k, while the amplitude and phase are determined by the starting position and velocity.

The velocity and acceleration of a simple harmonic oscillator oscillate with the same frequency as the position but with shifted phases. The velocity is maximum for zero displacement, while the acceleration is in the opposite direction as the displacement.

The potential energy stored in a simple harmonic oscillator at position x is

$$U = \frac{1}{2}kx^2.$$

Damped harmonic oscillator



Dependence of the system behavior on the value of the damping ratio ζ.

In real oscillators friction, or **damping**, slows the motion of the system. In many vibrating systems the frictional force F_f can be modeled as being proportional to the velocity v of the object: $F_f = -cv$, where c is called the *viscous damping coefficient*.

Newton's second law for damped harmonic oscillators is then

$$F = -kx - c\frac{dx}{dt} = m\frac{d^2x}{dt^2}.$$

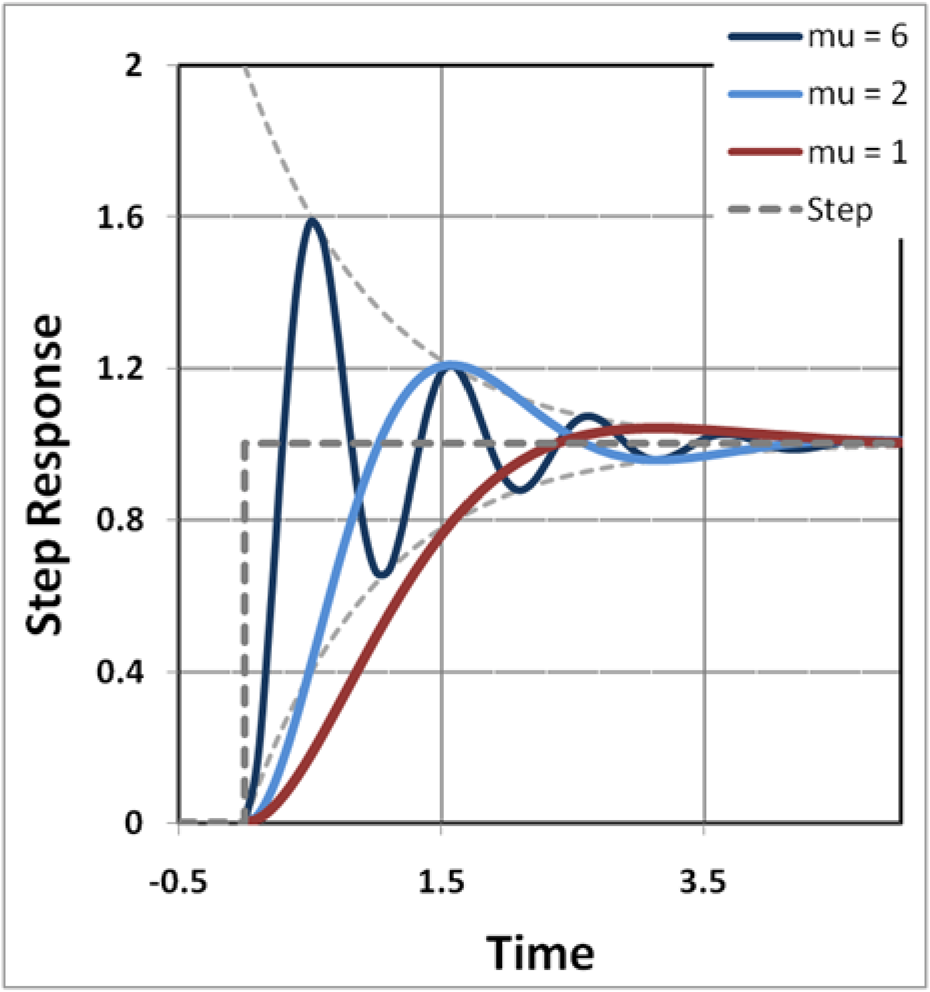
This is rewritten into the form

$$\frac{d^2x}{dt^2} + 2\zeta\omega_0\frac{dx}{dt} + \omega_0^2 x = 0,$$

where

$\omega_0 = \sqrt{\frac{k}{m}}$ is called the 'undamped angular frequency of the oscillator' and

$\zeta = \frac{c}{2m\omega_0}$ is called the 'damping ratio'.



Step-response of a damped harmonic oscillator; curves are plotted for three values of $\mu = \omega_1 = \omega_0\sqrt{1-\zeta^2}$. Time is in units of the decay time $\tau = 1/(\zeta\omega_0)$.

The value of the damping ratio ζ critically determines the behavior of the system. A damped harmonic oscillator can be:

- *Overdamped* ($\zeta > 1$): The system returns (exponentially decays) to equilibrium without oscillating. Larger values of the damping ratio ζ return to equilibrium slower.
- *Critically damped* ($\zeta = 1$): The system returns to equilibrium as quickly as possible without oscillating. This is often desired for the damping of systems such as doors.
- *Underdamped* ($\zeta < 1$): The system oscillates (with a slightly different frequency than the undamped case) with the amplitude gradually decreasing to zero. The angular frequency of the underdamped harmonic oscillator is given by

$$\omega_1 = \omega_0\sqrt{1 - \zeta^2}.$$

The Q factor of a damped oscillator is defined as

$$Q = 2\pi \times \frac{\text{Energy stored}}{\text{Energy lost per cycle}}.$$

Q is related to the damping ratio by the equation $Q = \frac{1}{2\zeta}.$

Driven harmonic oscillators

Driven harmonic oscillators are damped oscillators further affected by an externally applied force *F*(*t*).

Newton's second law takes the form

$$F(t) - kx - c\frac{dx}{dt} = m\frac{d^2x}{dt^2}.$$

It is usually rewritten into the form

$$\frac{d^2x}{dt^2} + 2\zeta\omega_0\frac{dx}{dt} + \omega_0^2 x = \frac{F(t)}{m} \ .$$

This equation can be solved exactly for any driving force using the solutions *z*(*t*) to the unforced equation, which satisfy:

$$\frac{d^2z}{dt^2} + 2\zeta\omega_0\frac{dz}{dt} + \omega_0^2 z = 0 \ ,$$

which can be expressed as a damped sinusoidal oscillation:

$$z(t) = Ae^{-\zeta\omega_0 t} \ \sin\left(\sqrt{1-\zeta^2}\ \omega_0 t + \varphi\right) \ ,$$

in the case where $\zeta \leq 1$. The amplitude *A* and phase φ determine the behavior needed to match the initial conditions.

Step input

In the case $\zeta < 1$ and a unit step input with $x(0) = 0$:

$$F(t) = \begin{cases} \omega_0^2 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

the solution is:

$$x(t) = 1 - e^{-\zeta\omega_0 t} \frac{\sin\left(\sqrt{1-\zeta^2}\,\omega_0 t + \varphi\right)}{\sin(\varphi)} \ ,$$

with phase φ given by

$$\cos\varphi = \zeta.$$

This behavior is found in (for example) feedback amplifiers, where the amplifier design is adjusted to obtain the fastest step response possible without undue overshoot or undershoot and with an adequate settling time.

The time an oscillator needs to adapt to changed external conditions is of the order $\tau=1/(\zeta\omega_0)$. In physics, the adaptation is called relaxation, and τ is called the relaxation time.

In electrical engineering, a multiple of τ is called the *settling time*, i.e. the time necessary to insure the signal is within a fixed departure from final value, typically within 10%. The term *overshoot* refers to the extent the maximum response exceeds final value, and *undershoot* refers to the extent the response falls below final value for times following the maximum response.

Sinusoidal driving force

In the case of a sinusoidal driving force:

$$\frac{d^2x}{dt^2} + 2\zeta\omega_0\frac{dx}{dt} + \omega_0^2 x = \frac{1}{m}F_0\sin(\omega t) \ ,$$

where $F_0$is the driving amplitude and ω is the driving frequency for a sinusoidal driving mechanism. This type of system appears in AC driven RLC circuits (resistor-inductor-capacitor) and driven spring systems having internal mechanical resistance or external air resistance.

The general solution is a sum of a transient solution that depends on initial conditions, and a steady state that is independent of initial conditions and depends only on the driving amplitude F_0, driving frequency, ω , undamped angular frequency ω_0, and the damping ratio ζ.

The steady-state solution is proportional to the driving force with an induced phase change of ϕ:

$$x(t) = \frac{F_0}{mZ_m\omega}\sin(\omega t + \phi)$$

where

$$Z_m = \sqrt{(2\omega_0\zeta)^2 + \frac{1}{\omega^2}\left(\omega_0^2 - \omega^2\right)^2}$$

is the absolute value of the impedance or linear response function and

$$\phi = \arctan\left(\frac{2\omega\omega_0\zeta}{\omega^2 - \omega_0^2}\right)$$

is the phase of the oscillation relative to the driving force.

For a particular driving frequency called the resonance frequency $\omega_r = \omega_0\sqrt{1 - 2\zeta^2}$, the amplitude (for a given F_0) is maximum. For underdamped systems the value of the amplitude can become quite large near the resonance frequency.

The transient solutions are the same as the unforced ($F_0 = 0$) damped harmonic oscillator and represent the systems response to other events that occurred previously. The transient solutions typically die out rapidly enough that they can be ignored.

Parametric oscillators

A parametric oscillator is a harmonic oscillator whose parameters oscillate in time. For example, a well known parametric oscillator is a child on a swing where periodically changing the child's center of gravity causes the swing to oscillate. The varying of the parameters drives the system. Examples of parameters that may be varied are its resonance frequency ω and damping β.

Parametric oscillators are used in many applications. The classical varactor parametric oscillator oscillates when the diode's capacitance is varied periodically. The circuit that varies the diode's capacitance is called the "pump" or "driver". In microwave electronics, waveguide/YAG based parametric oscillators operate in the same fashion. The designer varies a parameter periodically to induce oscillations.

Parametric oscillators have been developed as low-noise amplifiers, especially in the radio and microwave frequency range. Thermal noise is minimal, since a reactance (not a resistance) is varied. Another common use is frequency conversion, e.g., conversion from audio to radio frequencies. For example, the Optical parametric oscillator converts an input laser wave into two output waves of lower frequency (ω_s,ω_i).

Parametric resonance occurs in a mechanical system when a system is parametrically excited and oscillates at one of its resonant frequencies. Parametric excitation differs from forcing, since the action appears as a time varying modification on a system parameter. This effect is different from regular resonance because it exhibits the instability phenomenon.

Universal oscillator equation

The equation

$$\frac{\mathrm{d}^2q}{\mathrm{d}\tau^2} + 2\zeta\frac{\mathrm{d}q}{\mathrm{d}\tau} + q = 0$$

is known as the **universal oscillator equation** since all second order linear oscillatory systems can be reduced to this form. This is done through nondimensionalization.

If the forcing function is $f(t) = \cos(\omega t) = \cos(\omega t_c \tau) = \cos(\omega\tau)$, where $\omega = \omega t_c$, the equation becomes

$$\frac{\mathrm{d}^2q}{\mathrm{d}\tau^2} + 2\zeta\frac{\mathrm{d}q}{\mathrm{d}\tau} + q = \cos(\omega\tau).$$

The solution to this differential equation contains two parts, the "transient" and the "steady state".

Transient solution

The solution based on solving the ordinary differential equation is for arbitrary constants c_1 and c_2

$$q_t(\tau) = \begin{cases} e^{-\zeta\tau}\left(c_1 e^{\tau\sqrt{\zeta^2-1}} + c_2 e^{-\tau\sqrt{\zeta^2-1}}\right) & \zeta > 1 \text{ (overdamping)} \\ e^{-\zeta\tau}(c_1 + c_2\tau) = e^{-\tau}(c_1 + c_2\tau) & \zeta = 1 \text{ (critical damping)} \\ e^{-\zeta\tau}\left[c_1 \cos\left(\sqrt{1-\zeta^2}\tau\right) + c_2 \sin\left(\sqrt{1-\zeta^2}\tau\right)\right] & \zeta < 1 \text{ (underdamping)} \end{cases}$$

The transient solution is independent of the forcing function.

Steady-state solution

Apply the "complex variables method" by solving the auxiliary equation below and then finding the real part of its solution:

$$\frac{\mathrm{d}^2 q}{\mathrm{d}\tau^2} + 2\zeta\frac{\mathrm{d}q}{\mathrm{d}\tau} + q = \cos(\omega\tau) + i\sin(\omega\tau) = e^{i\omega\tau}.$$

Supposing the solution is of the form

$$q_s(\tau) = Ae^{i(\omega\tau+\phi)}.$$

Its derivatives from zero to 2nd order are

$$q_s = Ae^{i(\omega\tau+\phi)},\ \frac{\mathrm{d}q_s}{\mathrm{d}\tau} = i\omega Ae^{i(\omega\tau+\phi)},\ \frac{\mathrm{d}^2 q_s}{\mathrm{d}\tau^2} = -\omega^2 Ae^{i(\omega\tau+\phi)}.$$

Substituting these quantities into the differential equation gives

$$-\omega^2 Ae^{i(\omega\tau+\phi)} + 2\zeta i\omega Ae^{i(\omega\tau+\phi)} + Ae^{i(\omega\tau+\phi)} = (-\omega^2 A + 2\zeta i\omega A + A)e^{i(\omega\tau+\phi)} = e^{i\omega\tau}.$$

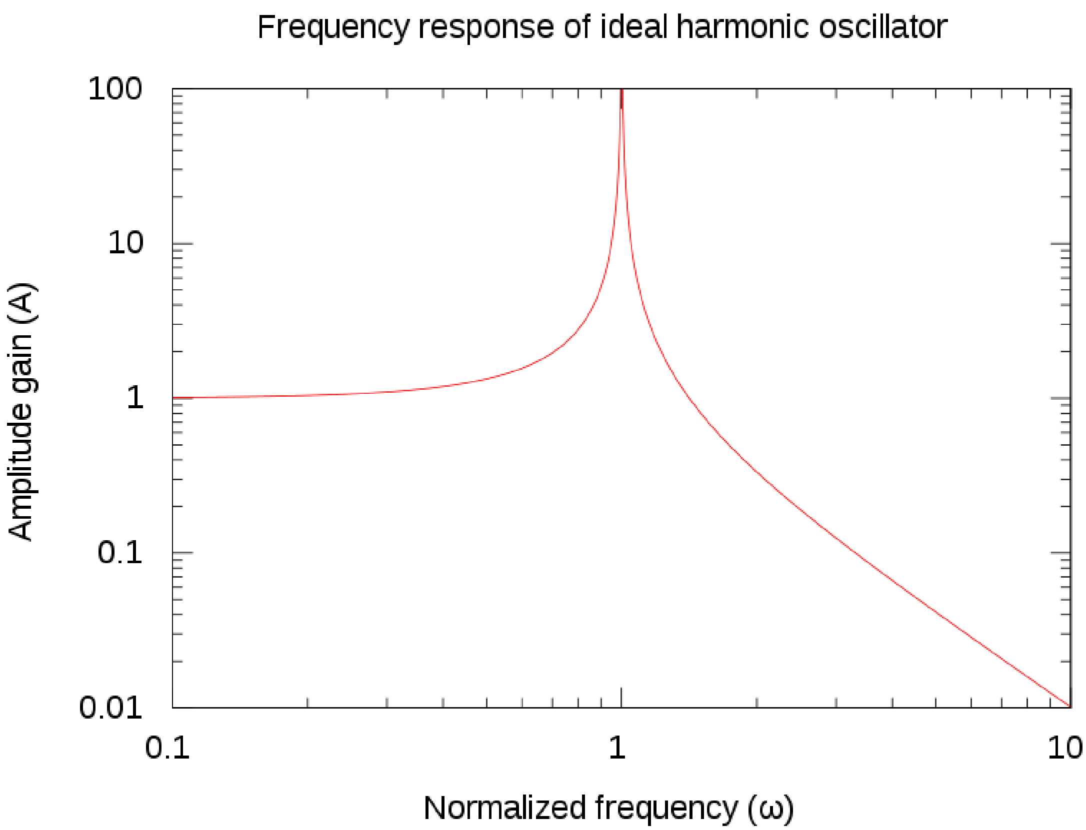
Dividing by the exponential term on the left results in

$$-\omega^2 A + 2\zeta i\omega A + A = e^{-i\phi} = \cos\phi - i\sin\phi.$$

Equating the real and imaginary parts results in two independent equations

$$A(1-\omega^2) = \cos\phi \qquad 2\zeta\omega A = -\sin\phi.$$

Amplitude part



Bode plot of the frequency response of an ideal harmonic oscillator.

Squaring both equations and adding them together gives

$$\left.\begin{aligned} A^2(1-\omega^2)^2 = \cos^2\phi \\ (2\zeta\omega A)^2 = \sin^2\phi \end{aligned}\right\} \Rightarrow A^2[(1-\omega^2)^2 + (2\zeta\omega)^2] = 1.$$

By convention the positive root is taken since amplitude is usually considered a positive quantity. Therefore,

$$A = A(\zeta,\omega) = \frac{1}{\sqrt{(1-\omega^2)^2 + (2\zeta\omega)^2}}.$$

Compare this result with the theory section on resonance, as well as the "magnitude part" of the RLC circuit. This amplitude function is particularly important in the analysis and understanding of the frequency response of second-order systems.

Phase part

To solve for φ, divide both equations to get

$$\tan\phi = -\frac{2\zeta\omega}{1-\omega^2} = \frac{2\zeta\omega}{\omega^2-1} \Rightarrow \phi \equiv \phi(\zeta,\omega) = \arctan\left(\frac{2\zeta\omega}{\omega^2-1}\right).$$

This phase function is particularly important in the analysis and understanding of the frequency response of second-order systems.

Full solution

Combining the amplitude and phase portions results in the steady-state solution

$$q_s(\tau) = A(\zeta,\omega)\cos(\omega\tau + \phi(\zeta,\omega)) = A\cos(\omega\tau+\phi).$$

The solution of original universal oscillator equation is a superposition (sum) of the transient and steady-state solutions

$$q(\tau) = q_t(\tau) + q_s(\tau).$$

Equivalent systems

Harmonic oscillators occurring in a number of areas of engineering are equivalent in the sense that their mathematical models are identical. Below is a table showing analogous quantities in four harmonic oscillator systems in mechanics and electronics. If analogous parameters on the same line in the table are given numerically equal values, the behavior of the oscillators—their output waveform, resonant frequency, damping factor, etc.—are the same.

Translational Mechanical	**Torsional Mechanical**	**Series RLC Circuit**	**Parallel RLC Circuit**
Position x	Angle θ	Charge q	Voltage e
Velocity $\frac{dx}{dt}$	Angular velocity $\frac{d\theta}{dt}$	Current $\frac{dq}{dt}$	$\frac{de}{dt}$
Mass M	Moment of inertia I	Inductance L	Capacitance C
Spring constant K	Torsion constant μ	Elastance $1/C$	Susceptance $1/L$

Friction γ	Rotational friction Γ	Resistance R	Conductance $1/R$
Drive force $F(t)$	Drive torque $\tau(t)$	e	di/dt
Undamped resonant frequency f_n:			
$\frac{1}{2\pi}\sqrt{\frac{K}{M}}$	$\frac{1}{2\pi}\sqrt{\frac{\mu}{I}}$	$\frac{1}{2\pi}\sqrt{\frac{1}{LC}}$	$\frac{1}{2\pi}\sqrt{\frac{1}{LC}}$
Differential equation:			
$M\ddot{x}+\gamma\dot{x}+Kx=F$	$I\ddot{\theta}+\Gamma\dot{\theta}+\mu\theta=\tau$	$L\ddot{q}+R\dot{q}+q/C=e$	$C\ddot{e}+\dot{e}/R+e/L=\dot{i}$

Application to a conservative force

The problem of the simple harmonic oscillator occurs frequently in physics, because a mass at equilibrium under the influence of any conservative force, in the limit of small motions, behaves as a simple harmonic oscillator.

A conservative force is one that has a potential energy function. The potential energy function of a harmonic oscillator is:

$$V(x)=\frac{1}{2}kx^2$$

Given an arbitrary potential energy function $V(x)$, one can do a Taylor expansion in terms of x around an energy minimum ($x = x_0$) to model the behavior of small perturbations from equilibrium.

$$V(x)=V(x_0)+(x-x_0)V'(x_0)+\frac{1}{2}(x-x_0)^2V^{(2)}(x_0)+O(x-x_0)^3$$

Because $V(x_0)$ is a minimum, the first derivative evaluated at x_0 must be zero, so the linear term drops out:

$$V(x)=V(x_0)+\frac{1}{2}(x-x_0)^2V^{(2)}(x_0)+O(x-x_0)^3$$

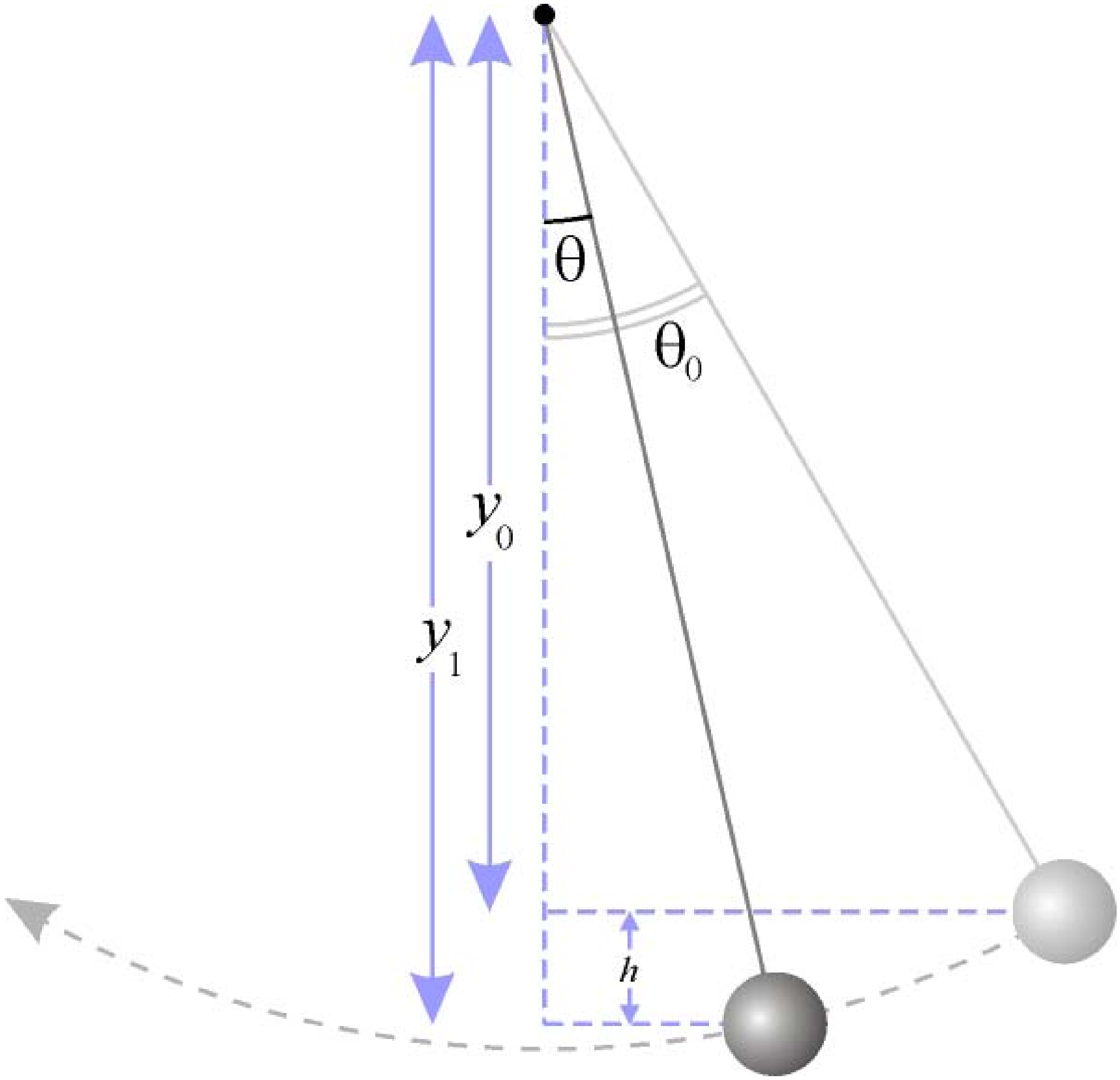
The constant term $V(x_0)$ is arbitrary and thus may be dropped, and a coordinate transformation allows the form of the simple harmonic oscillator to be retrieved:

$$V(x)\approx\frac{1}{2}x^2V^{(2)}(0)=\frac{1}{2}kx^2$$

Thus, given an arbitrary potential energy function $V(x)$ with a non-vanishing second derivative, one can use the solution to the simple harmonic oscillator to provide an approximate solution for small perturbations around the equilibrium point.

Examples

Simple pendulum



A simple pendulum exhibits simple harmonic motion under the conditions of no damping and small amplitude.

Assuming no damping and small amplitudes, the differential equation governing a simple pendulum is

$$\frac{d^2\theta}{dt^2} + \frac{g}{\ell}\theta = 0.$$

The solution to this equation is given by:

$$\theta(t) = \theta_0 \cos\left(\sqrt{\frac{g}{\ell}}t\right) \qquad\qquad |\theta_0| \ll 1$$

where θ_0 is the largest angle attained by the pendulum. The period, the time for one complete oscillation, is given by 2π divided by whatever is multiplying the time in the argument of the cosine ($\sqrt{\frac{g}{\ell}}$ here).

$$T_0 = 2\pi\sqrt{\frac{\ell}{g}} \qquad\qquad |\theta_0| \ll 1.$$

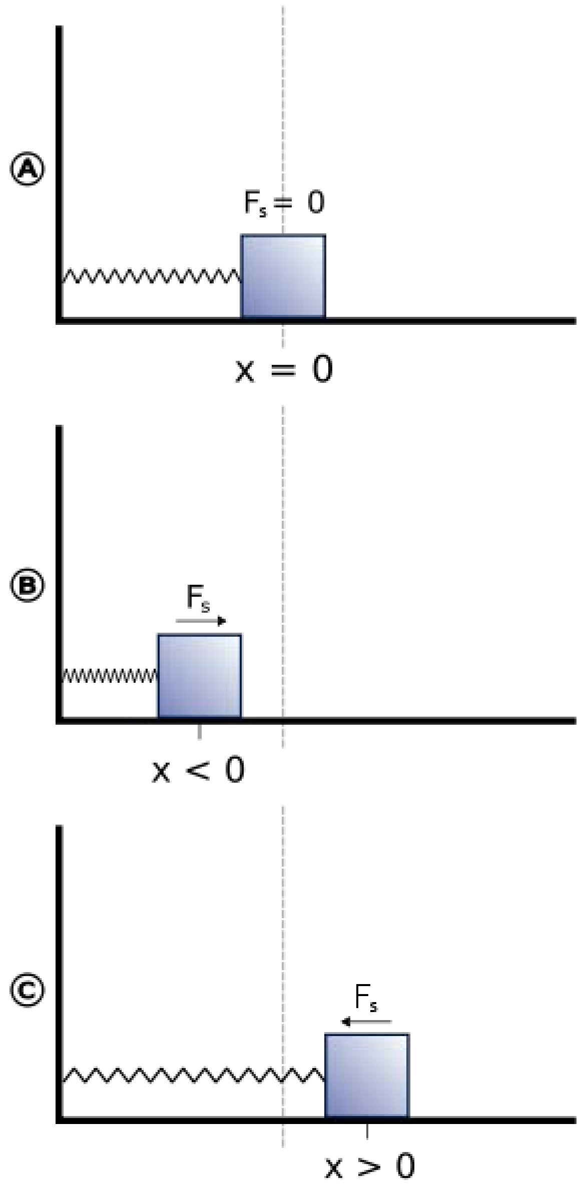
Pendulum swinging over turntable

Simple harmonic motion can in some cases be considered to be the one-dimensional projection of two-dimensional circular motion. Consider a long pendulum swinging over the turntable of a record player. On the edge of the turntable there is an object. If the object is viewed from the same level as the turntable, a projection of the motion of the object seems to be moving backwards and forwards on a straight line. When the frequency of rotation of the turntable has perfect synchronization with the motion of the pendulum, the angular speed of the turntable is known as the pulsation of the pendulum. Note that the pulsation of the pendulum is not the angular speed of the pendulum itself, but it is the angular speed of the corresponding circular motion.

In general, the pulsation of straight-line simple harmonic motion is the angular speed of the corresponding circular motion. Mathematically, a motion with period T and frequency $f=1/T$ has pulsation

$$\omega = 2\pi f = \frac{2\pi}{T}.$$

Spring–mass system



Spring–mass system in equilibrium (A), compressed (B) and stretched (C) states.

When a spring is stretched or compressed by a mass, the spring develops a restoring force. Hooke's law gives the relationship of the force exerted by the spring when the spring is compressed or stretched a certain length:

$$F(t) = -kx(t)$$

where F is the force, k is the spring constant, and x is the displacement of the mass with respect to the equilibrium position. This relationship shows that the distance of the spring is always opposite to the force of the spring.

By using either force balance or an energy method, it can be readily shown that the motion of this system is given by the following differential equation:

$$F(t) = -kx(t) = m\frac{\mathrm{d}^2}{\mathrm{d}t^2}x\left(t\right) = ma.$$

...the latter evidently being Newton's second law of motion.

If the initial displacement is A, and there is no initial velocity, the solution of this equation is given by:

$$x\left(t\right) = A\cos\left(\sqrt{\frac{k}{m}}t\right).$$

Given an ideal massless spring, *m* is the mass on the end of the spring. If the spring itself has mass, its effective mass must be included in *m*.

Energy variation in the spring–damping system

In terms of energy, all systems have two types of energy, potential energy and kinetic energy. When a spring is stretched or compressed, it stores elastic potential energy, which then is transferred into kinetic energy. The potential energy within a spring is determined by the equation $U = kx^2 / 2$.

When the spring is stretched or compressed, kinetic energy of the mass gets converted into potential energy of the spring. By conservation of energy, assuming the datum is defined at the equilibrium position, when the spring reaches its maximum potential energy, the kinetic energy of the mass is zero. When the spring is released, it tries to return to equilibrium, and all its potential energy converts to kinetic energy of the mass.